Social animal models for quantifying plasticity, assortment, and selection on interacting phenotypes

Martin, Jordan S ; Jaeggi, Adrian V

Abstract: Both assortment and plasticity can facilitate social evolution, as each may generate heritable associations between the phenotypes and fitness of individuals and their social partners. However, it currently remains difficult to empirically disentangle these distinct mechanisms in the wild, particularly for complex and environmentally responsive phenotypes subject to measurement error. To address this challenge, we extend the widely used animal model to facilitate unbiased estimation of plasticity, assortment, and selection on social traits, for both phenotypic and quantitative genetic analysis. Our social animal models (SAMs) estimate key evolutionary parameters for the latent reaction norms underlying repeatable patterns of phenotypic interaction across social environments. As a consequence of this approach, SAMs avoid inferential biases caused by various forms of measurement error in the raw phenotypic associations between social partners. We conducted a simulation study to demonstrate the application of SAMs and investigate their performance for both phenotypic and quantitative genetic analyses. With sufficient repeated measurements, we found desirably high power, low bias, and low uncertainty across model parameters using modest sample and effect sizes, leading to robust predictions of selection and adaptation. Our results suggest that SAMs will readily enhance social evolutionary research on a variety of phenotypes in the wild. We provide detailed coding tutorials and worked examples for implementing SAMs in the Stan statistical programming language.

DOI: https://doi.org/10.1111/jeb.13900

Social animal models for quantifying plasticity, assortment, and selection on interacting phenotypes

J. S. Martin[1], A. V. Jaeggi[1]

[1]Human Ecology Group, Institute of Evolutionary Medicine, University of Zurich, Switzerland

Corresponding author details:
Jordan S. Martin, email: jordan.martin@uzh.ch, ORCID: 0000-0001-8704-6076

Coauthor details:
Adrian V. Jaeggi, adrian.jaeggi@uzh.ch, ORCID: 0000-0003-1695-0388

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/JEB.13900

Abstract

Both assortment and plasticity can facilitate social evolution, as each may generate heritable associations between the phenotypes and fitness of individuals and their social partners. However, it currently remains difficult to empirically disentangle these distinct mechanisms in the wild, particularly for complex and environmentally responsive phenotypes subject to measurement error. To address this challenge, we extend the widely used animal model to facilitate unbiased estimation of plasticity, assortment, and selection on social traits, for both phenotypic and quantitative genetic analysis. Our social animal models (SAMs) estimate key evolutionary parameters for the latent reaction norms underlying repeatable patterns of phenotypic interaction across social environments. As a consequence of this approach, SAMs avoid inferential biases caused by various forms of measurement error in the raw phenotypic associations between social partners. We conducted a simulation study to demonstrate the application of SAMs and investigate their performance for both phenotypic and quantitative genetic analyses. With sufficient repeated measurements, we found desirably high power, low bias, and low uncertainty across model parameters using modest sample and effect sizes, leading to robust predictions of selection and adaptation. Our results suggest that SAMs will readily enhance social evolutionary research on a variety of phenotypes in the wild. We provide detailed coding tutorials and worked examples for implementing SAMs in the Stan statistical programming language.

Keywords: reaction norm, assortment, plasticity, social evolution, animal model

Introduction

Social interactions are central to the adaptive evolution of many complex phenotypes (Bourke, 2011). Sexual cooperation and competition, for example, can select for highly elaborated weapons, ornaments, and signals (Hare & Simmons, 2019; McCullough, Miller, & Emlen, 2016; Smith & Harper, 2003), as well as for novel mating systems and costly reproductive strategies (Díaz-Muñoz, DuVal, Krakauer, & Lacey, 2014; Hughes, Oldroyd, Beekman, & Ratnieks, 2008). The evolutionary ecology of social interactions has, therefore, been extensively investigated over the last half-century, employing both formal models and comparative empirical research across a diverse range of taxa (Bourke, 2011; Frank, 1998; Marshall, 2015; Rubenstein & Abbot, 2017). This work demonstrated the importance of social interactions as determinants of fitness (Cally, Stuart-Fox, & Holman, 2019; Frank, 2007; West et al., 2015), the fundamental roles of **social plasticity**, **assortment**, and **social selection** (see **Table 1**) in the evolutionary response to social interactions (Araya-Ajoy, Westneat, & Wright, 2020; Hamilton, 1964; Marshall, 2015; McGlothlin et al., 2014; Queller, 2011), as well as the ubiquity of phenotypic associations among social partners in wild populations (Brask et al., 2019; Carter et al., 2015; Janicke et al., 2019; Jiang, Bolnick, & Kirkpatrick, 2013).

Assortative mating is a particularly well studied form of assortment that may occur for a variety of plastic and multivariate phenotypes. For instance, Steller's Jay (Cyanocitta stelleri) breeding partners exhibit similar trait values across a behavioral syndrome of multiple exploratory and risk-taking behaviors, and pairs with more similar trait values have a higher probability of fledgling success (Gabriel & Black, 2012). Previous meta-analyses suggest that assortative mating may be widespread in animal populations, as evidenced by the ubiquity of positive associations between mating partners' phenotypes (Jiang et al., 2013). However, various alternative mechanisms may also cause these phenotypic associations to occur even in the absence of assortment, such as plasticity toward social partners, spatiotemporal

heterogeneity in the environment, and/or measurement error (Class et al., 2017; Wang et al., 2019).

Effectively distinguishing phenotypic associations caused by social plasticity from those caused by assortment per se is particularly crucial because each of these mechanisms may independently facilitate social evolution in the absence of the other (Araya-Ajoy et al., 2020; Marshall, 2015; McGlothlin et al., 2010). As Hamilton (1964) demonstrated, assortment potentiates a social evolutionary response by generating associations between individuals' genetic trait values and the fitness of their social partners (Bijma & Wade, 2008; McGlothlin et al., 2014; Queller, 2011). However, even when individuals interact randomly, social plasticity can still facilitate evolutionary change, as plastic trait values are determined not only by direct genetic effects on individual phenotypes, but also by indirect genetic effects (IGEs) due to heritable variation in the phenotypes of social partners. As a consequence, the social environment can also evolve whenever direct genetic effects on individual trait values are also associated with IGEs on the trait values or fitness of social partners (Bijma & Wade, 2008; Bijma, 2011).

Recent empirical studies have highlighted the importance of social plasticity and attendant IGEs across a diverse range of species, as well as the role of IGEs in potentiating evolutionary change (e.g. Bailey, Marie-Orleach, & Moore, 2017; Chenoweth, Rundle, & Blows, 2010; Evans et al., 2018; Santostefano et al., 2017; Silva et al., 2013; Wade et al., 2010). Both assortment and plasticity are, therefore, central for determining the evolutionary response to social interactions (McGlothlin et al., 2010). However, it remains difficult to disentangle the distinct effects of social plasticity, assortment, and measurement error in empirical datasets, as well as to integrate these mechanisms with information on the genetic causes and fitness consequences of measured phenotypes. Ultimately, this inhibits our ability to explain the causes of adaptive social evolution in the wild.

Evolutionary quantitative genetics has addressed this challenge with a powerful suite of theory for investigating the social evolution of interacting phenotypes (Araya-Ajoy et al., 2020; Bijma & Wade, 2008; Dingemanse & Araya-Ajoy, 2015; McGlothlin et al., 2010; Moore, Brodie, & Wolf, 1997; Wolf, Brodie, & Moore, 1999). Unfortunately, however, it remains difficult to avoid various sources of statistical and inferential bias when attempting to estimate these models in wild populations. In the present study, we therefore developed a series of novel quantitative genetic models called social animal models (SAMs), which can be used with repeated measurements data to distinguish the genetic and environmental effects of assortment, social plasticity, and both social and non-social selection on interacting phenotypes. These models are extensions of well-established animal models, which provide a generalized mixed-effects modelling framework for estimating the evolutionary quantitative genetic parameters of plastic traits (Nussey, Wilson, & Brommer, 2007; Wilson et al., 2010; de Villemereuil et al., 2016). Animal models are particularly important in evolutionary ecology because they facilitate inference of the individual **reaction norms** (**RNs**) underlying raw phenotypic measurements (Nussey et al., 2007). RNs are functions composed of individual-specific parameters, such as intercepts and slopes, that predict an individual's repeatable trait expression in response to an environmental factor, independently of other causes of phenotypic (co)variation. These RN parameters can be conceptualized as **intrinsic trait values** capturing individuals' differential patterns of phenotypic consistency (RN intercepts) and plasticity (RN slopes) across environments. Distinguishing the fitness effects of individuals' RN parameters is thus crucial for disentangling the evolutionary consequences of the environmentally responsive or unresponsive components of measured phenotypes (Nussey et al., 2007; Dingemanse et al., 2010). The SAMs presented here extend this basic animal model framework to appropriately estimate **social reaction norms** (SRNs) for interacting phenotypes, and thus

to achieve estimation of plasticity, assortment, and selection for individual differences in SRNs, independently of any nonrepeatable or unmeasured causes of phenotypic (co)variation.

We begin below by formally introducing the animal model and RNs for plastic traits, as well the general motivation for estimating animal models within a Bayesian framework using repeated measures data. We then present novel SAMs to address three key statistical challenges in the application of animal models to interacting phenotypes (**Figure 1**): (i) estimation of SRNs capturing feedback between the intrinsic trait values of individuals and their social partners, (ii) distinguishing the effects of assortment and plasticity on SRNs, and (iii) estimating selection and the response to selection caused by the SRN parameters of individuals and their social partners. A glossary of major conceptual terms is provided in **Table 1**, and a notation key is provided in **Table 2**. To further investigate the statistical properties of the proposed models, we conducted a simulation study to assess bias, uncertainty, and power for key evolutionary parameters with moderate sample and effect sizes. Our findings clearly demonstrate the utility of SAMs for both phenotypic and quantitative genetic analysis in empirical contexts comparable to many long-term field studies. Supplementary coding tutorials and worked examples of SAMs in the Stan statistical programming language (Carpenter et al., 2017) are also provided (see **Appendix S1**).

133 **Table 1** Glossary.

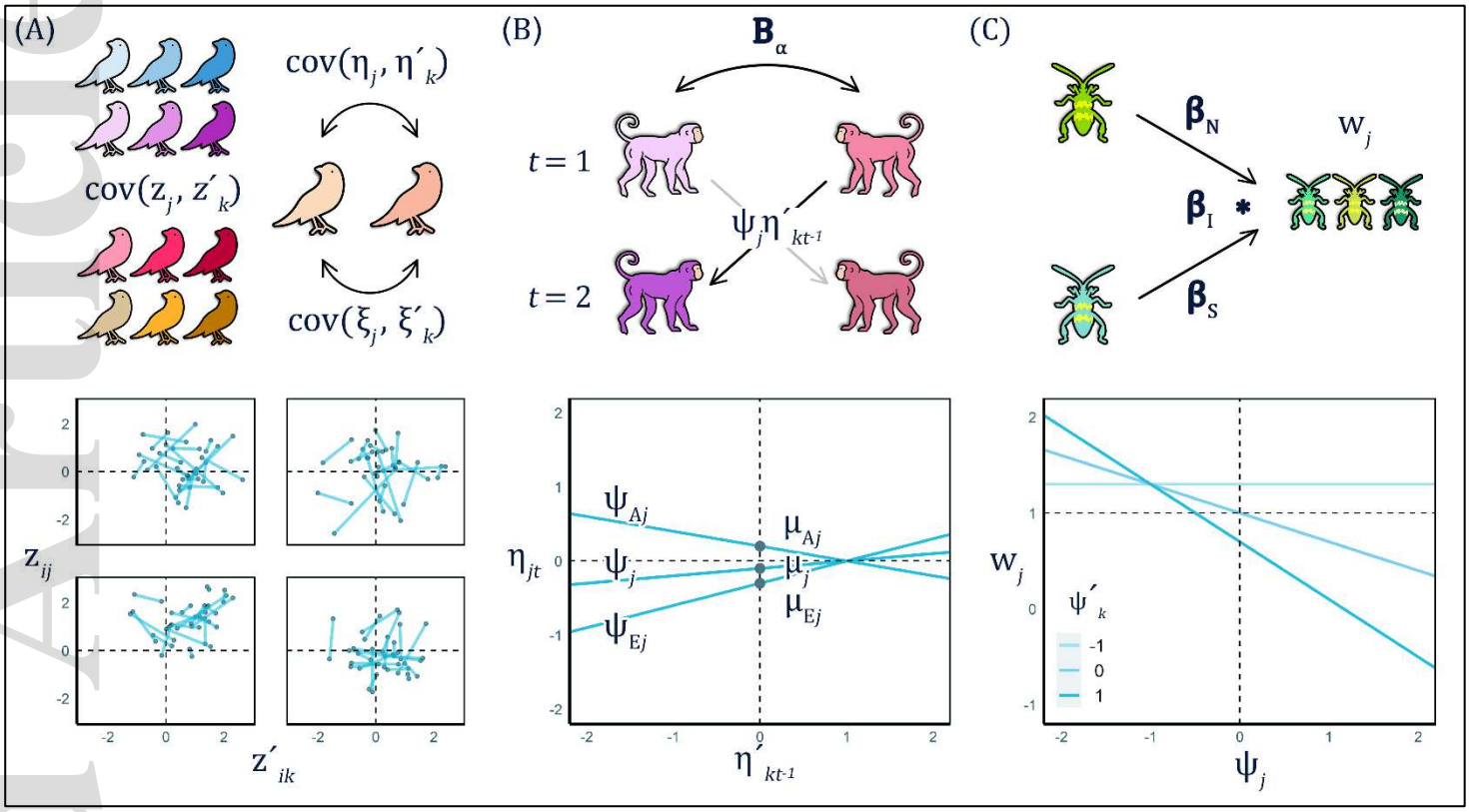
Term	Description
Assortment	The association between an individual's intrinsic trait value and the intrinsic trait value of their social partner(s), independent of any other causes of association between social partners' raw trait values (**Eq 4**).
Social plasticity	Phenotypic change in a focal individual caused by the traits of social partners, also referred to as social responsiveness (**Eq 2-3**). When partner phenotypes are heritable, social plasticity causes indirect genetic effects (IGEs) on the phenotype of the focal individual. Selection on social plasticity can, therefore, affect the magnitude of IGEs within a population (**Eq 9-10**).
Social selection	A systematic association between the intrinsic trait values of social partners and individual fitness in a population, due both to direct effects of partner phenotypes on individual fitness, as well as interactive effects caused by the joint trait values of individual and partner phenotypes (**Eq 5**).
Reaction norm	A reaction norm (RN) is a function predicting how an individual's phenotype will change in response to an environmental factor, independently of any non-repeatable causes of phenotypic change (**Eq 1**).
Social reaction norm	A social reaction norm (SRN) is a function predicting how an individual's phenotype will change in response to the phenotype of social partners (**Eq 3**).
Intrinsic trait value	A trait value that is solely attributable to direct and repeatable causes of between-individual variation, such as additive genetic and permanent environmental effects, but not indirect or non-repeatable within-individual effects, such as interactions with social partners. Throughout the paper, (S)RN parameters are defined as intrinsic trait values (**Eq 1** & **Eq 3**) subject to selection and adaptation (**Eq 5-8**).
SRN trait value	A trait value that is solely attributable to the SRN parameters of focal individuals and their interaction with the SRN parameters of social partners (**Eq 2-3**).

134

135 **Table 2.** Notation key**.**

Symbol	Meaning
i, j, k, t, A, E	Index of observation (i), focal individual (j), social partner (k), time point (t), additive genetic values (A), and permanent environmental values (E)
$\mathbf{z}, \mathbf{z}'$	Raw phenotypic measurements of focal individuals ($\boldsymbol{z}$) and their social partner(s) ($\boldsymbol{z}'$)
$\boldsymbol{\mu}$, $\boldsymbol{\mu}'$	Intrinsic trait values of an (S)RN intercept parameter for individuals ($\boldsymbol{\mu}$) and their social partners ($\boldsymbol{\mu}'$)
$\boldsymbol{\beta}$, $\boldsymbol{\beta}'$	Intrinsic trait values of an RN slope parameter for individuals ($\boldsymbol{\beta}$) and their social partners ($\boldsymbol{\beta}'$)
$\boldsymbol{\psi}$, $\boldsymbol{\psi}'$	Intrinsic trait values of an SRN slope parameter for individuals ($\boldsymbol{\psi}$) and their social partners ($\boldsymbol{\psi}'$), often referred to as interaction coefficients
$\boldsymbol{\eta}, \boldsymbol{\eta}'$	SRN trait values of modelled phenotypes for individuals ($\boldsymbol{\eta}$) and their social partners ($\boldsymbol{\eta}'$), which capture the direct and indirect (i.e. social) effects of SRN parameters
$\boldsymbol{\epsilon}, \boldsymbol{\epsilon}'$	Residual trait values of focal individuals ($\boldsymbol{\epsilon}$) and their social partners ($\boldsymbol{\epsilon}'$), capturing the phenotypic values explained by unmodelled and/or non-repeatable effects
$\boldsymbol{\xi}, \boldsymbol{\xi}', \phi$	SRN measurement error for individuals ($\boldsymbol{\xi}$) and their social partners ($\boldsymbol{\xi}'$), defined by residual trait values as well as any unmodelled causes of residual feedback (ϕ) across time
$\mathbf{P}, \mathbf{G}, \mathbf{E}$	Phenotypic (**P**), additive genetic (**G**) and permanent environmental (**E**) covariance matrices for (S)RN parameters
$\boldsymbol{\Sigma}$	Covariance matrix for residual trait values
β_α, $\mathbf{B}_\alpha$	The assortment coefficient (β_α) for the intrinsic trait value of individuals and their social partners, and the assortment matrix $\mathbf{B}_\alpha$ generalizing the assortment coefficient to multiple intrinsic trait values, such as SRN intercepts and slopes
$\boldsymbol{\beta}_\mathbf{N}$	Non-social selection gradients for SRN intercepts and slopes
$\boldsymbol{\beta}_\mathbf{S}$	Social selection gradients for SRN intercepts and slopes
$\boldsymbol{\beta}_\mathbf{I}$	Interaction coefficients for selection on SRN intercepts and slopes
$s_{\bar{\mu}}$, $s_{\bar{\psi}}$, $s_{\bar{\eta}}$	Selection differentials for the population SRN intercept ($s_{\bar{\mu}}$), SRN slope ($s_{\bar{\psi}}$), and SRN trait value ($s_{\bar{\eta}}$)
$\Delta\bar{\mu}$, $\Delta\bar{\psi}$, $\Delta\bar{\eta}$	Responses to selection for the population SRN intercept ($\Delta_{\bar{\mu}}$), SRN slope ($\Delta_{\bar{\psi}}$), and SRN trait value ($\Delta\bar{\eta}$)

Fig 1. Statistical challenges in the study of interacting phenotypes.



Caption. Each panel describes an inferential issue addressed by SAMs, with a heuristic representation above and accompanying data visualization below. Multiple clades commonly used in social evolutionary research (birds, primates, and beetles) are represented to demonstrate the diversity of systems to which SAMs can be applied. See **Appendix S1** for further details on the data simulation.

(A) Raw measurements confound social effects attributable to individuals' SRN trait values $\text{cov}(\eta_j, \eta_k')$ with residual effects attributable to SRN measurement error $\text{cov}(\xi_j, \xi'_k)$, such as spatiotemporal heterogeneity and/or interactions caused by unmeasured traits. This makes it difficult to reliably infer the direction and magnitude of social effects from the covariance of partners' observed phenotypes alone $\text{cov}(z_j, z'_k)$. To demonstrate this, the bottom panel shows four simulated individuals (each grid) interacting with twenty distinct social partners across two measurement periods (connected by each line). Although the population is characterized by positive assortment and social plasticity, positive slopes are not reliably observed between partners' trait values across dyads. This bias results from negatively associated residual effects across measurement periods, including residual feedback caused by unmeasured traits, as well as differences in intrinsic trait values between individuals and their social partners.

(**B**) Partners' phenotypes may covary because of assortment between individuals, as described by the assortment matrix $\mathbf{B}_{\alpha}$, or because of plasticity within individuals over time, as described by the SRN slope ψ_j for an individual at time t in response to their partner's SRN trait value η'_{kt-1} during the previous time interval t-1. Partitioning these distinct mechanisms is necessary to unbiasedly estimate individual differences in SRN intercepts μ_j and SRN slopes ψ_j. As is shown in the bottom panel, these SRN parameters can be further partitioned in underlying additive genetic (A) and permanent environmental trait values (E), which may differ both in magnitude and direction.

(**C**) Individual differences in SRN intercepts and slopes may have distinct effects on fitness, but these outcomes are confounded in a selection analysis of raw phenotypic measures. Selection can instead be modelled directly on individual-specific SRN parameters to investigate the multivariate evolution of the SRN function. An individuals' SRN parameters may have a direct influence on their own fitness ($\boldsymbol{\beta}_{\mathbf{N}}$), as may the SRN parameters of their social partners ($\boldsymbol{\beta}_{\mathbf{S}}$). Synergism or antagonism may also occur between the SRN parameters of individuals and their social partners,

leading to non-additive fitness effects (**$\boldsymbol{\beta}_{\mathbf{I}}$**). For illustrative purposes, the bottom panel shows the relative fitness of an individual w_j as a function of their own SRN slope and its interaction with the SRN slope of their partner. Although lower slopes are adaptive when the expected social partner exhibits an average (0) or high (+1) slope, this fitness advantage disappears when the social partner has a relatively low (-1) slope. In this case, $\beta_N = \beta_S = \beta_I = -0.3$.

Animal Models

Herein, we focus attention on basic animal models with Gaussian responses to simplify notation. A linear animal model can be specified for some observation i of phenotype z, as expressed by individual j in response to an environmental factor x_{ij}, such that

$$z_{ij} = \mu_0 + \mu_j + (\beta_1 + \beta_j)x_{ij} + \epsilon_{ij} \quad (1)$$

$$\mu_j = \mu_{Aj} + \mu_{Ej}, \qquad \beta_j = \beta_{Aj} + \beta_{Ej}$$

$$[\boldsymbol{\mu}_{\mathbf{A}}, \boldsymbol{\beta}_{\mathbf{A}}]^{\intercal} \sim \text{MVNormal}(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}): \mathbf{G} = \begin{bmatrix} \text{Var}(\boldsymbol{\mu}_{\mathbf{A}}) & \text{Cov}(\boldsymbol{\mu}_{\mathbf{A}}, \boldsymbol{\beta}_{\mathbf{A}}) \\ \text{Cov}(\boldsymbol{\beta}_{\mathbf{A}}, \boldsymbol{\mu}_{\mathbf{A}}) & \text{Var}(\boldsymbol{\beta}_{\mathbf{A}}) \end{bmatrix}$$

$$[\boldsymbol{\mu}_{\mathbf{E}}, \boldsymbol{\beta}_{\mathbf{E}}]^{\intercal} \sim \text{MVNormal}(\mathbf{0}, \mathbf{E} \otimes \mathbf{I}): \mathbf{E} = \begin{bmatrix} \text{Var}(\boldsymbol{\mu}_{\mathbf{E}}) & \text{Cov}(\boldsymbol{\mu}_{\mathbf{E}}, \boldsymbol{\beta}_{\mathbf{E}}) \\ \text{Cov}(\boldsymbol{\beta}_{\mathbf{E}}, \boldsymbol{\mu}_{\mathbf{E}}) & \text{Var}(\boldsymbol{\beta}_{\mathbf{E}}) \end{bmatrix}$$

$$\boldsymbol{\epsilon} \sim \text{Normal}(0, \boldsymbol{\Sigma}): \boldsymbol{\Sigma} = [\text{Var}(\boldsymbol{\epsilon})]$$

Bold symbols are used to distinguish vectors and matrices from scalars, and the ᵀ symbol indicates the transpose operation. This animal model decomposes population-, individual- and observation-level effects on the measured phenotype. On average across the population, phenotypic expression is best predicted by the fixed global intercept μ_0 and regression coefficient β_1 representing plasticity toward the observed environmental factor x_{ij}. However, individuals also consistently differ in their patterns of phenotypic expression, such that responses are further affected by individual-specific RN intercepts μ_j and slopes β_j, represented here as random deviations from population-level values μ_0 and β_1. Note that repeated individual measurements are necessary to empirically partition these parameters.

When quantitative genetic information is available, these phenotypic RN parameters can be further partitioned into the sum of their underlying additive genetic (μ_{Aj}, β_{Aj}) and permanent environmental (μ_{Ej}, β_{Ej}) values. The phenotypic association between individuals' RN intercepts and slopes can then be expressed as the sum of distinct genetic and permanent environmental (co)variance matrices **G** and **E**, respectively. Assuming the genetic and environmental effects are independently distributed in the population, the total phenotypic covariance of RN parameters can be given by $\mathbf{P} = \mathbf{G} + \mathbf{E}$. Genetic covariance is scaled by the Kronecker product $\otimes$ of **G** with a relatedness matrix **A** derived from pedigree or molecular data, such that individuals with higher genetic similarity are expected to have more similar $\boldsymbol{\mu}_{\mathbf{A}}$ and $\boldsymbol{\beta}_{\mathbf{A}}$ values. Permanent environmental covariance is scaled by an identity matrix **I**, indicating that values are independent and identically distributed among individuals within the respective vectors $\boldsymbol{\mu}_{\mathbf{E}}$ and $\boldsymbol{\beta}_{\mathbf{E}}$. If factors such as spatiotemporal heterogeneity or maternal effects cause covariance among individuals' permanent environmental effects (e.g. Heckerman et al., 2016; Kruuk & Hadfield, 2007), a matrix capturing common environmental effects can instead be utilized (see Thomson et al. 2018 for suggestions). In the absence of genetic information, **G** and **E** cannot be distinguished and the animal model reduces to the standard linear mixed-effects model commonly used for phenotypic analysis of **P** in evolutionary ecology (Dingemanse & Dochtermann, 2013; Nussey et al., 2007). Finally, in addition to the deterministic effects of the individual and population parameters, each measurement is further subject to stochastic effects $\boldsymbol{\epsilon}$ caused by unmeasured factors uncorrelated with the other linear predictors. These residual values are randomly distributed with a variance described by the $\boldsymbol{\Sigma}$ matrix. Without repeated measurements, the permanent environmental effects **E** and non-permanent, residual effects $\boldsymbol{\Sigma}$ are confounded.

Benefits of the animal model

As previously noted, the central benefit of the animal model is its ability to distinguish individuals' RNs and intrinsic trait values from the various residual effects that also influence raw trait values. For labile phenotypes that are repeatedly expressed, individual measurements tend to exhibit modest repeatability across time (e.g. Bell, Hankison, & Laskowski, 2009; Cauchoix et al., 2018; Fanson & Biro, 2019). As a consequence, raw measurements often provide more information about unobserved environmental heterogeneity than about individuals' intrinsic trait values, thus confounding distinct causes of between- and within-individual (co)variation (Searle, 1961; Nakagawa & Schielzeth, 2010; Dingemanse & Dochtermann, 2013).

In addition to its methodological benefits, the animal model also provides deeper theoretical insight into the relationship between raw phenotypic measures and the intrinsic trait values that are ultimately subject to natural selection. Given that individuals can vary both in their RN intercepts and slopes, raw phenotypic observation z_{ij} is not merely an error-prone observation of a single individual trait value z_j. Rather, each raw measurement is an error-prone composite of distinct intrinsic trait values for the RN parameters governing individual differences in phenotypic consistency (intercept μ_j) and phenotypic plasticity (slope β_j) across the modelled environmental factor. These intrinsic trait values may be caused by separable sources of genetic and environmental (co)variation, as indicated by their corresponding parameters in the **G** and **E** matrices, and may thus experience distinct selection pressures (e.g. Ramakers, Gienapp, & Visser, 2019; Weis & Gorman, 1990). Therefore, animal models are important not only for partitioning RNs from raw phenotypic data, but also for understanding how the differential fitness effects of responsive (RN slopes) and non-responsive (RN intercepts) phenotypic components may lead to distinct patterns of evolutionary change.

Bayesian animal models

Standard animal models are often estimated and interpreted within a classical statistical framework, such as with the ASReml program (Gilmour et al., 2002). However, Bayesian estimation of animal models using Markov Chain Monte Carlo (MCMC) can also be readily implemented in the R package MCMCglmm (Hadfield, 2010) or the Stan statistical programming language (Carpenter et al., 2017). The linear animal model in **Eq 1**, for example, can be estimated within a Bayesian framework by specifying prior distributions for the unknown model parameters

$$\mu_0, \beta_1, \mathbf{G}, \mathbf{E}, \mathbf{\Sigma} \sim \boldsymbol{f}_{\text{prior}}\left(\mathbf{\Theta}_{\text{prior}}\right)$$

This notation indicates that the model parameters have prior distributions characterized by probability density functions $\boldsymbol{f}_{\text{prior}}$ with parameters $\mathbf{\Theta}_{\text{prior}}$, such as the general purpose $\text{Normal}(0,1)$ or $\text{Half}-\text{Cauchy}(0,1)$ priors recommended by Lemoine (2019).

While there are many benefits to Bayesian inference in general (see McElreath, 2020 for a detailed treatment), Bayesian animal models are particularly useful because of their ability to quantify uncertainty in individuals' RN parameters, as well as to carry this uncertainty forward across multiple stages of analysis (Hadfield et al., 2010; Stinchcombe, Simonsen, & Blows, 2014; Martin 2021). Individual-specific intercept and slope values are often estimated with high degrees of uncertainty, particularly for small sample sizes and traits with moderate to low repeatability. Conducting subsequent analyses with point estimates of these values, also known as best linear unbiased predictors (BLUPs), leads to a undesirable risk of inferential bias and anti-conservative inference (Hadfield et al., 2010; Houslay & Wilson, 2017). However, in the Bayesian animal model, individual RNs are no longer estimated with single expected values $\hat{\mu}_j$ and $\hat{\beta}_j$; instead, parameters are characterized by posterior distributions $\Pr(\mu_j | \mathbf{z}, \mathbf{\Theta})$ and $\Pr(\beta_j | \mathbf{z}, \mathbf{\Theta})$ that fully capture the probabilistic uncertainty in these estimates,

conditional on the observed responses $\mathbf{z}$ and other model parameters and priors $\boldsymbol{\Theta}$. When estimated with Markov Chain Monte Carlo (MCMC), parameters are approximated by vectors of posterior samples that can be used to calculate new quantities of interest, facilitating straightforward statistical inference for values that are not directly specified as parameters in the model. As is explored further below, flexible Bayesian modelling software such as Stan (Carpenter et al., 2017) can also specify multi-stage analyses within a single model, simultaneously accounting for uncertainty in the estimation of RNs and their effects on fitness or other phenotypes (Martin, 2021).

Social Animal Models

We now extend the basic animal model (**Eq 1**) to account for the effects of social interactions on phenotypic expression and fitness, which presents a series of unique statistical challenges (**Figure 1**). As noted above, SAMs address these challenges by (i) partitioning SRNs for intrinsic trait values from other non-repeatable or unmeasured causes of variation, (ii) distinguishing the effects of assortment and plasticity on associations among social partners, and (iii) estimating selection and the response to selection caused by individual variation in SRN intercepts and slopes. These models are based on an extensive body of prior theory for studying the evolutionary quantitative genetics of interacting phenotypes. We therefore address each challenge (i-iii) in a stepwise fashion, building up SAMs sequentially to better identify their relation to previous theoretical models of IGEs, as well as to highlight important empirical considerations.

IGE models account for the indirect effects of the social environment on individual phenotypes (Bijma, 2011; Bijma & Wade, 2008; McAdam, Garant, & Wilson, 2014; McGlothlin et al., 2010; Moore et al., 1997; Wolf et al., 1999). So-called variance-partitioning and trait-based IGE models provide two distinct but potentially equivalent parameterizations

for empirical research (McGlothlin & Brodie, 2009; Bijma, 2014), each with their own benefits and drawbacks. The variance-partitioning approach describes the total (co)variance attributable to the effects of individuals and their social partners on trait expression and fitness (Bijma, 2011; Bijma & Wade, 2008). As a consequence, this approach facilitates accurate predictions of evolutionary change (Morrissey et al., 2010; Morrissey et al., 2012; Price, 1972; Robertson, 1966), but it can also obscure the direct and indirect causal pathways underlying selection on these variance components (Hadfield & Thomson, 2017). In contrast, the trait-based approach considers the specific phenotypes causing social effects on trait expression and fitness (McGlothlin et al., 2010; Moore et al., 1997; Wolf et al., 1999), thus distinguishing between the direct and indirect effects of distinct traits. This parameterization is crucial for effectively testing adaptive hypotheses of integrated phenotypes (Lande & Arnold, 1983; McGlothlin et al., 2010). However, trait-based models will also be biased by the exclusion of relevant phenotypes or forms of phenotypic interaction, making them sensitive to misspecification and biased predictions of evolutionary change (Bijma, 2014; Morrisey et al., 2012). The SAMs presented here are based primarily on trait-based IGE models, as this approach is able to distinguish between social plasticity, assortment, and direct and indirect selection on specific phenotypes. However, these approaches can always be integrated by including additional variance components in the models presented below (Dingemanse & Araya-Ajoy, 2015; McAdam, Garant, & Wilson, 2014).

i. Estimating social reaction norms

Classic trait-based IGE models do not account for individual variation in social plasticity, instead assuming that plasticity is a fixed trait in the population and thus does not undergo selection (Moore et al., 1997; Wolf et al., 1999). However, intraspecific variation in social plasticity has been found across a variety of systems, suggesting that many phenotypes may be better described by SRNs. For example, individual differences in social information

use have been observed across diverse taxa ranging from humans (Molleman, van der Berg, & Weissing, 2014) and chimpanzees (Pan troglodytes; Watson et al., 2018) to water dragons (*Intellegama lesueurii*; Strickland & Frère, 2019) and scops owls (Otus scops; Parejo & Avilés, 2020), such that some individuals are consistently more responsive to the behavior of their social partners than others. The evolution of social plasticity in sexual display traits has also been demonstrated experimentally in fruit flies (Drosophila serrata; Chenoweth et al., 2010). The presence of variable social plasticity within a population provides the opportunity for selection on both the environmentally responsive and unresponsive components of a phenotype, which can further affect the rate and direction of evolutionary change (Araya-Ajoy et al., 2020; Kazancıoğlu, Klug, & Alonzo, 2012; McGlothlin et al., 2021; McNamara & Leimar, 2020; Van Cleve, 2017; Van Cleve & Akçay, 2014). The fitness benefits of social plasticity may, for instance, be frequency-dependent, causing selection to maintain individual variation in responsiveness toward social partners (Wolf, van Doorn, & Weissing, 2008). In such cases, it is critical to accurately estimate SRNs in order to best explain differential patterns of selection across social environments, as well as to avoid any inferential biases caused by unmodelled causes of association between the phenotypes of social partners.

Measurement error of SRNs in trait-based models

Kazancıoğlu et al. (2012) investigated the evolutionary consequences of SRNs using trait-based IGE models with heritable variation in SRN intercepts and slopes. Consider a simple linear interaction between the phenotype z_j of individual j and the phenotype z'_k of social partner k, with primes $'$ used herein to denote the values of social partners. This social interaction can be modelled such that

$$z_j = \mu_j + \psi_j z'_k \tag{2.1}$$

$$z'_k = \mu'_k + \psi'_k z_j$$

The SRN slope ψ_j is often referred to as an interaction coefficient and quantifies the social plasticity of the focal individual in response to their social partner's phenotype. Feedback will occur when partners express social plasticity in the same phenotypes over time, with each individual sequentially increasing or decreasing their phenotypic expression in response to the trait value of the partner. We can see this by substituting in the phenotype z'_k of the social partner into the focal phenotype z_j and simplifying (Moore et al., 1997). Focusing on the response of the focal individual

$$z_j = \mu_j + \psi_j z'_k = \frac{\mu_j + \psi_j \mu'_k}{1 - \psi_j \psi'_k} \quad (2.2)$$

The trait value z_j is thus a function of the individual's SRN intercept μ_j, their SRN slope in response to the SRN intercept of the social partner $\psi_j \mu'_k$, as well as feedback due to the interaction of SRN slopes between the individual and their social partner $1 - \psi_j \psi'_k$. When either partner is not socially responsive, i.e. $\psi_j \psi'_k = 0$, then the feedback effect is removed.

Despite animal models being well suited for investigating non-social RNs, it is difficult to empirically estimate SRNs and their feedback effects without inferential bias, as associations between partners' phenotypes may be caused by a variety of distinct mechanisms (**Figure 1A**; Class et al., 2017; Wang et al., 2019). For the remainder of the paper, we use the term 'measurement error' in a broad but formal sense to refer to any variation that produces a difference between estimated or observed values and the true values of a trait, with specific attention toward the interaction of individuals' SRN intercepts and slopes on measured traits. From this perspective, any effects that cause raw measurements to deviate from the repeatable trait values predicted by SRN interactions can be considered measurement error with respect to those values, consistent with the use of this term in the statistical literature (Bollen & Noble, 2011; Loken & Gelman, 2017). Estimating assortment and social plasticity on the repeatable

components of measured trait interactions, independently of measurement error, is central for social evolutionary analysis, as it is only the associations between social partners' intrinsic trait values that will contribute to evolutionary change (Araya-Ajoy et al., 2020; Bijma, 2011).

The influence of SRN measurement error on statistical inferences can be seen by introducing an additional vector of phenotypic residuals $\boldsymbol{\epsilon}$ into the formal model (**Eq 2.2**)

$$z_j = \mu_j + \psi_j z'_k + \epsilon_j = \frac{\mu_j + \psi_j \mu'_k}{1 - \psi_j \psi'_k} + \frac{\epsilon_j + \psi_j \epsilon'_k}{1 - \psi_j \psi'_k} \quad (2.3)$$

To enhance clarity, we can further distinguish between the **SRN trait value** (η_j), which is caused by the interaction of the focal and social partners' intrinsic trait values for SRN parameters, and the measurement error with respect to the SRN trait values (ξ_j), such that

$$z_j = \eta_j + \xi_j \quad (2.4)$$

$$\eta_j = \frac{\mu_j + \psi_j \mu'_k}{1 - \psi_j \psi'_k}, \qquad \xi_j = \frac{\epsilon_j + \psi_j \epsilon'_k}{1 - \psi_j \psi'_k}$$

The SRN measurement errors of focal individuals $\boldsymbol{\xi}$ and their social partners $\boldsymbol{\xi}'$ are by definition independent of their respective SRN trait values $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$, so that the covariance between an individual and their social partner's raw trait values is simply

$$\mathrm{cov}(z_j, z'_k) = \mathrm{cov}(\eta_j, \eta'_k) + \mathrm{cov}(\xi_j, \xi'_k) \quad (2.5)$$

The phenotypic (co)variance of raw measurements $\mathrm{cov}(z_j, z'_k)$ is therefore attributable both to repeatable covariance caused by SRN trait values $\mathrm{cov}(\eta_j, \eta'_k)$ as well as any sources of shared SRN measurement $\mathrm{cov}(\xi_j, \xi'_k)$ (**Figure 1A**).

Inferential bias caused by SRN measurement error

This trait-based IGE model shows how the use of raw trait values can lead to inferential bias when applying standard animal models to interactions among labile phenotypes. Firstly, note that as a consequence of feedback in **Eq 2**, individuals' raw trait values (z_j, z'_k) are expected to associate with the residual trait values of their respective social partners (ϵ'_k, ϵ_j), such that $\mathrm{cov}(z_j, \epsilon'_k) \neq 0$ and $\mathrm{cov}(z'_k, \epsilon_j) \neq 0$. However, standard animal models assume that residuals are statistically independent of predictor variables, including the trait values of social partners. As a consequence, empirical estimates of SRN slopes obtained from a trait-based animal model will tend to be biased, particularly for small to moderately sized slopes (see Bijma, 2014 for further discussion). This so-called endogeneity bias can be avoided with variance-partitioning approaches (e.g. Bijma, 2014; Koster et al., 2015). However, there is currently no general solution for avoiding endogeneity bias with trait-based models.

A distinct but related source of bias is the assumption that social plasticity of constant magnitude is the only cause of association between partners' phenotypes, including residual covariance between raw trait values. This assumption is reflected in **Eq 2** by the use of the same trait-specific SRN slope ψ_j to scale the covariance due to intrinsic and residual trait values, which is a consequence of defining these parameters on raw measurements $\boldsymbol{z}$ and $\boldsymbol{z}'$. While it may seem sensible to estimate the magnitude of social plasticity using observed phenotypes, various unmeasured effects of differing magnitude can also cause raw phenotypic associations between social partners over short and long timescales (Westneat, Wright, & Dingemanse, 2015; Class et al., 2017; Wang et al., 2019). In addition to factors such as spatiotemporal heterogeneity and researcher bias, these residual associations can also reflect social effects caused by other unmeasured traits, which may be subject to distinct magnitudes of social plasticity (i.e. distinct trait-specific SRN slopes). Such residual effects tend to be much

larger than the repeatable component of measured phenotypes in both laboratory and field settings, particularly for labile traits such as behavior, cognitive performance, or hormone levels (Bell, Hankison, & Laskowski, 2009; Cauchoix et al., 2018; Fanson & Biro, 2019). As a consequence, standard models using raw trait values to calculate $\mathrm{cov}(z_j, z'_k)$ will tend to bias the magnitude of SRN effects for measured traits, i.e. $\mathrm{cov}(\eta_j, \eta'_k)$, with the magnitude of residual effects, i.e. $\mathrm{cov}(\xi_j, \xi'_k)$ (Brommer, 2013; Dingemanse & Dochtermann, 2013).

Two central challenges in extending the animal model to interacting phenotypes are, therefore, to effectively avoid endogeneity bias caused by feedback, as well as to avoid inferential bias caused by various sources of SRN measurement error. Fortunately, we can draw on the flexibility of Bayesian animal models to specify SAMs that explicitly partition SRN and residual effects during model estimation, avoiding inferential bias caused by using raw trait values subject to measurement error. Herein we discuss two SAMs for repeated measurements within or across social partners. We then introduce SAMs for repeated measures both within and across partners, along with a general solution for partitioning the effects of assortment and social plasticity in these models.

SAMs for repeated measures within partners

Although it is convenient to formally model the long-run expectation of feedback between social partners, such as in the standard SRN trait model (**Eq 2.4)**, empirical datasets often contain interactions of heterogeneous duration and may include repeated measurements within interactions. We therefore need to differentiate the consequences of social interactions between specific measurement periods, in addition to the more basic challenge of differentiating the effects of intrinsic and residual trait values. In particular, we can use a time index t to indicate the measurement period within a particular social interaction, e.g. $t = \{1,2,3,4\}$ for an individual measured four times while interacting with the same partner. With

this longitudinal information, a so-called autoregressive moving average (ARMA) function, commonly used for time-series analysis, can then be implemented to differentiate the effects of feedback and stochastic variation across measurement periods (Box, Jenkins, Reinsel, & Ljung, 2016). The basic idea of ARMAs is to regress the trait value at time t on the trait value(s) at a previous time such as t-1 to account for "autoregressive" feedback effects, i.e. $z_t \sim z_{t-1}$, as well as on the residual trait value(s) at time t-1 to account for any unmeasured factors causing the average response to "move" between sampling periods, i.e. $z_t \sim \epsilon_{t-1}$.

To further differentiate repeatable and residual feedback due to SRN parameters, as well as to avoid endogeneity bias, we propose an extension of the basic ARMA function to further separate social feedback on the latent SRN parameters from all other residual effects on measured trait values causing SRN measurement error. This allows estimation of SRN slopes independently of the magnitude of residual covariation among partners, thus relaxing the assumptions of classical trait-based models (**Eq 2**). In particular, for both phenotypic and quantitative genetic analysis, we propose the following SAM for repeated measurements of focal and social partner phenotypes

$$\mathrm{z}_{jt} = \mu_0 + \eta_{jt} + \xi_{jt} \tag{3.1}$$

$$\eta_{jt} = \begin{cases} \mu_j + (\psi_1 + \psi_j)\mu'_{\mathrm{k}} & \text{if } t = 1 \\ \mu_j + (\psi_1 + \psi_j)\eta'_{kt-1} & \text{else} \end{cases}$$

$$\xi_{jt} = \begin{cases} \epsilon_{jt} & \text{if } t = 1 \\ \epsilon_{jt} + \phi\epsilon'_{kt-1} & \text{else} \end{cases}$$

$$\mu_j = \mu_{\mathrm{A}j} + \mu_{\mathrm{E}j}, \quad \psi_j = \psi_{\mathrm{A}j} + \psi_{\mathrm{E}j}$$

$$[\boldsymbol{\mu}_{\mathbf{A}}, \boldsymbol{\mu}'_{\mathbf{A}}, \boldsymbol{\psi}_{\mathbf{A}}, \boldsymbol{\psi}'_{\mathbf{A}}]^{\mathrm{T}} \sim \mathrm{MVNormal}(\mathbf{0}, \mathbf{G} \otimes \mathbf{A})$$

$$[\boldsymbol{\mu}_{\mathbf{E}}, \boldsymbol{\mu}'_{\mathbf{E}}, \boldsymbol{\psi}_{\mathbf{E}}, \boldsymbol{\psi}'_{\mathbf{E}}]^{\mathrm{T}} \sim \mathrm{MVNormal}(\mathbf{0}, \mathbf{E} \otimes \mathbf{I})$$

$$[\boldsymbol{\epsilon}, \boldsymbol{\epsilon}']^{\mathrm{T}} \sim \mathrm{Normal}(0, \boldsymbol{\Sigma}): \boldsymbol{\Sigma} = \begin{bmatrix} \mathrm{Var}(\boldsymbol{\epsilon}) & \mathrm{Cov}(\boldsymbol{\epsilon}, \boldsymbol{\epsilon}') \\ \mathrm{Cov}(\boldsymbol{\epsilon}', \boldsymbol{\epsilon}) & \mathrm{Var}(\boldsymbol{\epsilon}') \end{bmatrix}$$

$$\mu_0, \psi_1, \phi, \mathbf{G}, \mathbf{E}, \boldsymbol{\Sigma} \sim \boldsymbol{f}_{\mathrm{prior}}(\boldsymbol{\Theta}_{\mathrm{prior}})$$

To address the sequential structure of social interactions, the ARMA function specifies distinct feedback processes between the latent SRN and residual trait values of individuals and their social partners. The SRN trait value η_{jt} plastically responds to the partner SRN trait value η'_{kt-1} over time as a function of the individual-specific interaction coefficient ψ_j; similarly, the SRN measurement error ξ_{jt} captures any unmeasured feedback effects caused by the residual trait values of social partners at a previous time ϵ'_{kt-1}, which are independent of the repeatable SRN effects of z'_{kt} on z_{jt} (and vice versa), with a distinct residual feedback coefficient ϕ, as well as the remaining individual residual ϵ_{jt}. Note that the partner response model is defined equivalently with respect to the focal individual and is thus not shown explicitly for brevity. However, separate model parameters can also be accommodated when the responses of focal individuals and their social partners systematically differ, such as when sexes exhibit distinct patterns of social plasticity (e.g. Strickland & Frère, 2019). See **Appendix S1** for further details on implementing a variety of extensions to this basic SAM.

By directly partitioning the distinct effects of SRN and residual feedback, rather than confounding them as in **Eq 2**, the estimation of SRN trait values, and in particular the SRN slopes, is expected to be unbiased by the magnitude of (co)variance attributable to SRN measurement error. As a consequence, this SAM facilitates estimation of the magnitude of social plasticity and IGEs on the intrinsic trait values of any social phenotype, irrespective of other sources of residual (co)variation between measurements, including phenotypic interactions caused by unmeasured traits. Accounting for and distinguishing both sources of temporal feedback in the model also removes the risk of endogeneity bias in SRN parameters.

Note that the population intercept μ_0 is intentionally separated from the ARMA process, as accumulation of the mean during feedback may lead to unrealistic predictions, particularly when the interaction coefficients are very large or small. However, there may be traits for which individuals' absolute rather than relative values are of primary interest (Westneat et al., 2020), in which case the model can be reparametrized appropriately. Higher-order and/or nonlinear ARMA effects can also be straightforwardly accommodated if animals exhibit more complex response surfaces (Box et al., 2016). Similarly, the ARMA process currently assumes that individuals are being measured in the context of an ongoing social interaction. The best unbiased prediction of the individual's SRN trait value is thus given by their SRN intercept and their SRN slope (social plasticity) toward the SRN intercept of their partner, i.e. $\eta_{jt=1} = \mu_j + \psi_j \mu_k'$, with subsequent temporal change modelled through the autoregressive feedback process. However, if individuals are instead measured prior to exposure to conspecifics, the function can simply be redefined such that $\eta_{jt=1} = \mu_j$ in the absence of social interaction.

SAMs for repeated measures across partners

In some social systems, it may be easier to gather repeated measurements across multiple social partners rather than within the same partner. Some sampling methods, such as observational sampling of behavior, may also require aggregation across many repeated measurements to achieve effective estimates of repeatable trait values (Koski, 2011). In such cases, temporal information on within-partner interactions will be missing, so that observations within each partner are effectively t = 1, and the index i is again utilized to distinguish repeated individual measurements across partners. In the absence of additional temporal information, the residual feedback effect ϕ cannot be directly partitioned and the SRN measurement errors $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ will reduce to the residual trait values $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}'$. A between-partner SAM can thus be specified by simply reducing the within-partner SAM (**Eq 3.1**) to ignore t > 1 feedback effects

$$z_{ijt=1} = \mu_0 + \eta_{ijt=1} + \epsilon_{ijt=1} \quad (3.2)$$

$$\eta_{ijt=1} = \mu_j + (\psi_1 + \psi_j)\mu_k'$$

The model priors and parameter distributions are otherwise equivalent to **Eq 3.1**, and the partner response model is again defined equivalently with respect to the focal individual. Although temporal effects are confounded in the residuals, this SAM still effectively partitions SRN trait values from SRN measurement error, as the SRN slopes are appropriately scaled by the latent SRN intercepts of social partners, rather than their raw trait values. As a consequence, neither residual covariance nor endogeneity are expected to bias inferences of SRN parameters from this model.

ii. Distinguishing SRN assortment and plasticity

Confounding of within- and between-individual (co)variation is a common source of bias in observational studies (van de Pol & Wright, 2009). Sprau and Dingemanse (2017), for instance, demonstrated that the association between risky behavior and urbanization across great tits (Parus major) reflects the tendency of bolder individuals to more frequently inhabit areas with high motor traffic, rather than a plastic response across individuals to motor traffic. Social plasticity and assortment can also be easily confounded when individuals are non-randomly distributed across social environments. In this case, associations among the intrinsic trait values of social partners may be caused by within-individual plasticity toward the trait value of the social partner, or by between-individual assortment caused by processes such as habitat selection or partner choice (**Figure 1B**). Despite being well recognized as a source of bias across non-social environmental factors, less attention has been given to confounding of social plasticity and assortment among interacting phenotypes, particularly in non-human social systems (cf. Steglich, Snijders, & Pearson, 2010).

SAMS for repeated measures within and between partners

Fortunately, however, a within-individual centering procedure (van de Pol & Wright, 2009) can be used to specify SAMs that effectively partition social plasticity and assortment whenever individuals are measured with multiple social partners. In particular, to isolate the appropriate SRN slopes **ψ** for the social plasticity of the focal individual, the ARMA function in **Eq 3.1** needs to be specified toward the deviation of each social partner's SRN trait value from the average SRN trait value experienced by the focal individual across social partners. In other words, each measurement of the social environment, as defined by the time-dependent SRN trait value η'_{ikt} of the current social partner, needs to be centered on the average SRN intercepts $\bar{\mu}'_K$, slopes $\bar{\psi}'_K$, and time-dependent trait value $\bar{\eta}'_{iKt}$ experienced by an individual across the set K of their social partners. We use the notation $\eta_{\mathrm{W}ijt}$ herein to indicate the within-individual centered SRN trait value for observation i of individual j at time t with respect to their current social partner. We also use $\eta_{\mathrm{B}ijt}$ to indicate the SRN trait value for the interaction of the focal individual and their average social partner, and we introduce an additional between-individual regression coefficient β_{B} to scale the average partner feedback process, which may reflect the effects of plasticity as well as assortment. Partitioning the within- and between-individual SRN trait values in this way appropriately adjusts the estimated SRN parameters for unbalanced sampling across partner trait values (van de Pol & Wright, 2009).

A SAM for repeated measures within and between partners can thus be specified by centering **Eq 3.1** within individuals such that

$$z_{ijt} = \mu_0 + \eta_{\mathrm{W}ijt} + \beta_{\mathrm{B}}\eta_{\mathrm{B}ijt} + \xi_{ijt} \tag{3.3}$$

$$\eta_{\mathrm{W}ijt} = \begin{cases} \mu_j + (\psi_1 + \psi_j)(\mu'_{\mathrm{k}} - \bar{\mu}'_K) & \text{if } t = 1 \\ \mu_j + (\psi_1 + \psi_j)(\eta'_{ikt-1} - \bar{\eta}'_{iKt-1}) & \text{else} \end{cases}$$

$$\eta_{\mathrm{B}ijt} = \begin{cases} (\psi_1 + \psi_j)\bar{\mu}'_K & \text{if } t = 1 \\ (\psi_1 + \psi_j)\bar{\eta}'_{iKt-1} & \text{else} \end{cases}$$

$$\xi_{ijt} = \begin{cases} \epsilon_{ijt} & \text{if } t = 1 \\ \epsilon_{ijt} + \phi\epsilon'_{ikt-1} & \text{else} \end{cases}$$

Centering of the partner SRN slope ψ'_{k} on the average partner SRN slope $\bar{\psi}'_K$ is specified implicitly through the definition of the partner η'_{ikt-1} and average partner $\bar{\eta}'_{iKt-1}$ SRN trait values. With the addition of β_{B}, all unspecified priors and generative distributions are otherwise equivalent to **Eq 3.1**, and the social partner response is defined equivalently with respect to the focal individual. Note that this within-individual centering procedure can also be used to distinguish plasticity and assortment in the between partner model (**Eq 3.2)** following the specification of $\eta_{\mathrm{W}ijt=1}$ and $\eta_{\mathrm{B}ijt=1}$ above. This SAM is based on a model previously proposed by Dingemanse and Araya-Ajoy (2015), but it avoids inferential bias caused by calculating within-individual deviations on raw trait values subject to SRN measurement error. Instead, social plasticity is modelled toward the individual-specific deviation of the SRN trait values, such that the between-individual effect of the average social environment does not confound the estimation of social plasticity.

Quantifying assortment across partners

This within-individual centered SAM accounts for the effects of non-randomly distributed social environments, but it does not directly parameterize assortment among social partners. Although the parameter β_{B} is necessary to reduce inferential bias in the SRN parameters, it does not provide a direct estimate of assortment necessary for predicting the response to social selection. Westneat et al. (2020) demonstrate that such between-individual regression coefficients can also be undesirably sensitive to model specification, in contrast to the more robust estimation of within-individual effects such as $\eta_{\mathrm{W}ijt}$. Therefore, β_{B} provides an unreliable means by which to estimate assortment on SRN parameters, and should instead

be conceptualized as a pragmatic parameter for reducing bias in the estimation of SRN parameters. Class et al. (2017) propose an alternative variance-partitioning approach for quantifying assortment using multi-response GLMMs with correlated SRN parameters across social partners. These models provide an important tool for reliably detecting assortment under realistic field conditions and will generate unbiased estimates of assortment whenever social plasticity and IGEs are absent. However, only trait-based IGE models such as the proposed SAMs can partition the effects of both SRN assortment and social plasticity, which is crucial for testing causal hypotheses of adaptive social evolution (Araya-Ajoy et al., 2020; Hadfield & Thomson, 2017).

Rather than attempting to parameterize assortment directly in the SAM, we can instead treat assortment as a generative property of the model estimated from the posterior distributions of individual-level parameters. As discussed above, Bayesian inference via MCMC facilitates carrying uncertainty forward across any quantity or analysis defined over the posterior distributions of a model (Hadfield et al., 2010; Stinchcombe et al., 2014). Therefore, assortment among any group of social partners can be estimated by the association between the posteriors of their intrinsic SRN parameter trait values. Following McDonald et al. (2017), we define the phenotypic assortment coefficient β_α as the simple regression coefficient of the mean partner phenotype on the individual phenotype. These assortment coefficients can be readily estimated for SRN intercepts $\beta_{\bar{\mu}'\mu}$ and slopes $\beta_{\bar{\psi}'\psi}$ using posterior distributions of a SAM

$$\beta_{\bar{\mu}'\mu} = \Pr\left(\frac{\mathrm{cov}(\boldsymbol{\mu}, \bar{\boldsymbol{\mu}}')}{\mathrm{var}(\boldsymbol{\mu})}\middle|\mathbf{z}, \mathbf{z}', \boldsymbol{\Theta}\right) \quad (4.1)$$

$$\beta_{\bar{\psi}'\psi} = \Pr\left(\frac{\mathrm{cov}(\boldsymbol{\psi}, \bar{\boldsymbol{\psi}}')}{\mathrm{var}(\boldsymbol{\psi})}\middle|\mathbf{z}, \mathbf{z}', \boldsymbol{\Theta}\right)$$

These regression coefficients will be equivalent to a Pearson correlation coefficient whenever variance is constant across individual and social partner phenotypes. Note that in

dyadic contexts without multiple partners, the average partner phenotypes will simply be the trait values of the social partner. Average partner phenotypes across multiple interactions can also be scaled to estimate the expected assortment within a single interaction (**Appendix S1**).

More generally, a matrix $\mathbf{B}_\alpha$ can be estimated to account for assortment between SRN intercepts and slopes, i.e. $\beta_{\bar{\mu}'\psi}$ and $\beta_{\bar{\psi}'\mu}$

$$\mathbf{B}_\alpha = \Pr\left(\begin{bmatrix} \beta_{\bar{\mu}'\mu} & \beta_{\bar{\psi}'\mu} \\ \beta_{\bar{\mu}'\psi} & \beta_{\bar{\psi}'\psi} \end{bmatrix} \middle| \mathbf{z}, \mathbf{z}', \boldsymbol{\Theta}\right) \quad (4.2)$$

Assortment is here assumed to be a fixed property of the population. However, the magnitude of assortment may also be an additional SRN parameter subject to variation and selection, in which case it would be appropriate to estimate individual- rather than population-level assortment coefficients across multiple selection events (see Araya-Ajoy et al., 2020 for further discussion).

By excluding the assortment coefficients from the model specification, the priors of the SAM assume that partner phenotypes are statistically independent, conditional on the relatedness matrix **A**, so that the prior probability of assortment is centered on $\mathbf{B}_\alpha = \mathbf{0}$. All else being equal, this conservative assumption will tend to regularize the posterior assortment coefficients toward null values. Nonetheless, with sufficiently informative datasets, the joint likelihood of the SAM will exert a much stronger influence on the shape and location of the individual posterior distributions, leading to accurate empirical estimates of assortment irrespective of the model priors. This approach allows for highly flexible estimation of assortment on any trait values of interest, among partners in groups of any size, without adding unnecessary complexity to the basic statistical model.

iii. Selection and the response to selection on SRNs

Rather than expressing selection on measured trait values, a fitness model can instead be specified for selection directly on SRN parameters, effectively distinguishing between selection on the responsive and non-responsive components of individual phenotypes (Dingemanse et al., 2010; Kazancıoğlu et al., 2012), as well as avoiding bias caused by residual effects on raw measurements. This selection analysis for SRN parameters can be readily accomplished with the SAM by adding an additional response model to **Eq 3** for predicting individual fitness w_j, such that

$$w_j = \nu_0 + \beta_{N1}\mu_j + \beta_{N2}\psi_j + \beta_{S1}\bar{\mu}'_K + \beta_{S2}\bar{\psi}'_K + \beta_{I1}\left(\mu_j\bar{\mu}'_K\right) + \beta_{I2}\left(\psi_j\bar{\psi}'_K\right) + \delta_j \quad (5)$$

$$\nu_0, \boldsymbol{\beta_N}, \boldsymbol{\beta_S}, \boldsymbol{\beta_I}, \mathrm{var}(\boldsymbol{\delta}) \sim \boldsymbol{f}_{\mathrm{prior}}\left(\boldsymbol{\Theta}_{\mathrm{prior}}\right)$$

Parameters and corresponding priors are specified for the population intercept of fitness υ_0, the non-social ($\boldsymbol{\beta_N}$) and social ($\boldsymbol{\beta_S}$) selection gradients, interactive ($\boldsymbol{\beta_I}$) coefficients on SRN intercepts and slopes, and the variance of any residual effects on fitness $\mathrm{var}(\boldsymbol{\delta})$. Note that mean partner trait values $\bar{\mu}'_K$ and $\bar{\psi}'_K$ are used to account for the possibility of multiple partners during a selection event, such as when a single lifetime fitness measure is available for individuals with multiple lifetime partners. This parameterization can also be further extended to account for the effects of larger social groups by scaling mean partner values with the expected number of partners per selection event $\bar{n}$ (McGlothlin et al., 2010). Whenever selection is instead estimated with a single social partner, such as within a breeding season for monogamous species, the partner trait values can be substituted for the mean partner values.

This fitness model builds on previous extensions of the Lande and Arnold (1983) framework to interacting phenotypes (Araya-Ajoy et al., 2020; Frank, 1998; Queller, 2011;

Westneat, 2012; Wolf et al., 1999), which did not consider the effects of SRN measurement error on evolutionary inference. In addition to the additive fitness effects of individual ($\boldsymbol{\beta_N}$) and partner SRN parameters ($\boldsymbol{\beta_S}$), synergistic or antagonistic effects may also occur between SRN parameters ($\boldsymbol{\beta_I}$), so that the payoffs of these trait values are contingent on the trait values of the social partner (**Figure 1C**). When biologically relevant, further interactive effects can also be added to the fitness model for the joint trait values of SRN intercepts and slopes, i.e. $\beta_{I3}(\mu_j \bar{\psi}'_K) + \beta_{I4}(\psi_j \bar{\mu}'_K)$, as well as for other nonlinear effects of interest. These interactive effects of joint trait values can be interpreted as the degree to which individual payoffs deviate from additivity across different social environments (Marshall, 2015; Queller, 2011). For example, in some ecological contexts, biparental care will lead to higher fitness payoffs for both sexes than uniparental care, so that the highest fitness is expected among individuals who both engage in offspring care and mate with partners who also engage in offspring care (Alger et al., 2020; Pilakouta, Hanlon, & Smiseth, 2018; Kokko & Johnstone, 2002). In Blackcaps (Sylvia atricapilla), who experience high rates of nest predation, similar degrees of care between parents have been found to result in faster rates of nestling growth (Leniowski & Węgrzyn, 2018). Similarly, burying beetles (Nicrophorus vespilloides) who cooperate in biparental care tend to rear larger offspring with a higher probability of survival to adulthood (Pilakouta et al., 2018). Please see Araya-Ajoy et al. (2020) for a deeper treatment of these interactive effects and their importance for social evolution.

To enhance interpretation, selection gradients are specified in the model for individuals' SRN parameter deviations**,** rather than on absolute SRN parameter values (i.e. population average + individual values), which centers the fitness model on the expected population values μ_0 and ψ_1. This transformation can be easily adjusted, however, if variation in population means is of biological interest, such as when studying the effects of frequency-dependent selection across multiple episodes of selection (Araya-Ajoy et al., 2020). The fitness model

further assumes that selection gradients are equivalent across individuals, so that the fitness function is symmetric between individual j and individual k. This assumption can be relaxed by introducing distinct fitness models for multiple classes of individuals (e.g. males and females), as described above for **Eq 3**. Selection coefficients estimated from non-Gaussian SAMs can also be transformed to appropriate selection gradients for evolutionary prediction following the approach of Morrissey and Sakrejda (2013). Finally, note that the SAMs of phenotypic expression and fitness should generally be estimated together in Stan as a single multi-response model (i.e. **Eq. 3** + **Eq. 5**), as uncertainty in the estimation of SRN parameters will thereby be represented simultaneously in both trait models. This ability to specify key evolutionary parameters simultaneously across response models is a central benefit of the proposed modelling framework, as it avoids a variety of issues caused by alternative multi-stage, "stats-on-stats" approaches (see Martin, 2021 for further discussion).

Estimating the response to selection

The SAM approach to selection analysis provides a straightforward means to calculate selection differentials directly for SRN parameters, which will represent the within-generation change in SRN parameters following an episode of selection (Lande, 1979). Assuming that SRN parameters are centered on zero and fitness is appropriately mean-scaled, **Eq 5** can be substituted into the Robertson-Price identity (Price, 1972; Robertson, 1966) to derive selection differentials (McGlothlin et al., 2010) for SRN intercepts $s_{\bar{\mu}}$ and slopes $s_{\bar{\psi}}$

$$\mathbf{s} = \begin{bmatrix} s_{\bar{\mu}} \\ s_{\bar{\psi}} \end{bmatrix} = \mathbf{P}\boldsymbol{\beta}_{\mathbf{N}} + \mathbf{C}\boldsymbol{\beta}_{\mathbf{S}} \quad (6)$$

Where $\mathbf{P} = \mathbf{G} + \mathbf{E}$ is a phenotypic (co)variance matrix of SRN parameters (**Eq. 3**) and $\mathbf{C}$ is a matrix of (co)variances among individuals' SRN parameters and the mean partner parameters experienced in the social environment. With all residual and indirect effects excluded from the

SRN parameters, the **C** matrix of partner covariances is solely attributable to covariance caused by assortment. In particular,

$$\mathbf{C} = \mathrm{diag}(\mathbf{P})\mathbf{B}_{\boldsymbol{\alpha}} \tag{7}$$

where $\mathrm{diag}(\mathbf{P})$ is a matrix with the variances of SRN parameters on the diagonal and $\mathbf{B}_{\boldsymbol{\alpha}}$ is the assortment matrix defined above (**Eq 4.2**). See **Appendix S1** for a detailed discussion and derivation of **Eq 6-7**. The genetic response to selection on SRN intercepts $\Delta\bar{\mu}$ and slopes $\Delta\bar{\psi}$ can then be estimated by substituting additive genetic effects **G** for the total phenotypic effects **P** in **s** (**Eq. 6**), partialling out the independent environmental effects **E** from the selection differential. This provides a multivariate breeder's equation (Lande & Arnold, 1983) of evolutionary change in SRN parameters

$$\begin{bmatrix}\Delta\bar{\mu} \\ \Delta\bar{\psi}\end{bmatrix} = \mathbf{G}\mathbf{P}^{-1}\mathbf{s} = \mathbf{G}\boldsymbol{\beta}_{\mathbf{N}} + \mathrm{diag}(\mathbf{G})\mathbf{B}_{\boldsymbol{\alpha}}\boldsymbol{\beta}_{\mathbf{S}} \tag{8}$$

These responses in SRN parameters can also be used to estimate the response in SRN trait values $\Delta\bar{\eta}$, which reflect both the direct effects and IGEs of SRN evolution. As noted by Kazancıoğlu et al. (2012), it is cumbersome to derive analytic solutions for the response in SRN trait values subject to feedback, and simulation using population parameters provides a clear alternative. However, solutions can be straightforwardly derived for the response in the absence of/prior to feedback, which can be tested empirically using both within- and between-partner sampling designs (**Eq 3.1-3.3**). In particular, the initial (t = 1) SRN trait value within generation (1) of a population in response to an average partner is given by

$$\bar{\eta}_{t=1}^{(1)} = \bar{\mu} + \bar{\psi}\bar{\mu}' + \mathrm{var}(\boldsymbol{\psi})\beta_{\bar{\mu}'\psi} \tag{9.1}$$

where $\mathrm{var}(\boldsymbol{\psi})\beta_{\bar{\mu}'\psi} = \mathrm{cov}(\boldsymbol{\psi}\bar{\boldsymbol{\mu}}')$ accounts for the effect of assortment. Assuming $\bar{\mu} = \bar{\mu}'$, $\bar{\psi} = \bar{\psi}'$, and the absence of selection on $\mathrm{var}(\boldsymbol{\psi})\beta_{\bar{\mu}'\psi}$, the responses in SRN parameters (**Eq 8**) can be substituted in for the expected SRN trait value at t = 1 in the subsequent generation (2)

$$\bar{\eta}_{t=1}^{(2)} = \bar{\mu} + \Delta\bar{\mu} + (\bar{\psi} + \Delta\bar{\psi})(\bar{\mu} + \Delta\bar{\mu}) + \mathrm{var}(\boldsymbol{\psi})\beta_{\bar{\mu}'\psi} \quad (9.2)$$

Subtracting **Eq 9.1** from **Eq 9.2** provides the response in the SRN trait value

$$\Delta\bar{\eta}_{t=1} = \Delta\bar{\mu} + \bar{\psi}\Delta\bar{\mu} + \Delta\bar{\psi}\bar{\mu} + \Delta\bar{\psi}\Delta\bar{\mu} \quad (9.3)$$

Where $\bar{\psi}\Delta\bar{\mu}$ is the change in IGEs expected in the absence of selection on SRN slopes, while $\Delta\bar{\psi}\bar{\mu}$ and $\Delta\bar{\psi}\Delta\bar{\mu}$ reflect further change in IGEs caused by selection on SRN slopes (Kazancıoğlu et al., 2012). Selection differentials for SRN trait values $s_{\bar{\eta}}$ can be similarly calculated by substituting in the SRN parameter selection differentials (**Eq 6**) such that

$$s_{\bar{\eta}_{t=1}} = s_{\bar{\mu}} + \bar{\psi}s_{\bar{\mu}} + s_{\bar{\psi}}\bar{\mu} + s_{\bar{\psi}}s_{\bar{\mu}} \quad (10)$$

The social evolution of SRNs can, therefore, be straightforwardly estimated using the proposed SAMs (**Eq 3** + **Eq 5**), with Bayesian inference providing additional information on the probabilistic uncertainty of these estimates (Stinchcombe et al., 2014). It should be noted that although **Eq 6-10** remove the biasing effect of SRN measurement error on evolutionary parameters, their predictions will nonetheless be sensitive to the exclusion of other fitness-relevant phenotypes, which is a general limitation of trait-based models of evolutionary change (Bijma, 2014; Morrissey et al., 2010). Further suggestions for interpreting and plotting the evolution of the broader SRN function within the population can be found in Martin (2021).

Simulation Study

General overview

We simulated empirical datasets to investigate the statistical properties of SAMs for phenotypic and quantitative genetic analysis at modest sample sizes typical of field research (N = 100, 200, 300). Data were simulated for aggressive interactions within a population of biannually breeding animals forming seasonal monogamous pairs (**Figure 2**), as is common in many avian taxa. There were always N/2 males and females, and each individual paired with four breeding partners across the study period. Aggression was measured twice in each subject during their social interactions with their breeding partner, resulting in eight repeated measures per individual across four breeding seasons. We therefore utilized the SAM defined in **Eq 3.3** for repeated measures within and between partners.

Observations within a breeding season were made using an experimental assay applied during initial (t = 1) and follow-up (t = 2) sampling periods. SRN feedback effects occurred across measurements due to aggressive interactions within a breeding season. SRN measurement error was also generated by residual feedback effects due to unmeasured factors, as well as other unspecified residual effects such as spatiotemporal heterogeneity that further caused partners' raw aggression measures to covary. To provide a direct demonstration of SAMs ability to differentiate covariance due to social plasticity and assortment, we further assumed that breeding partners assortatively mated for SRN slopes (i.e. more/less socially plastic birds tended to pair with more/less socially plastic partners). A single measure of reproductive success was taken for each pair at the end of the breeding season, resulting in four repeated measures of this fitness proxy per individual. While this dense sampling procedure will be unrealistic for some social systems and fitness components, previous simulations have shown that repeated sampling within pairs appreciably enhances statistical power for detecting

social effects (Class et al., 2017). Therefore, our simulation investigated what can be achieved at modest sample size with study designs prioritizing repeated individual measurement over the lifespan, as is common in many long-term field studies.

SAMs for both phenotypic and quantitative genetic analysis were estimated with these simulated datasets using Stan (Carpenter et al., 2017) in the R statistical environment (R Core Team, 2013). We simulated 200 datasets per sample size (N = 100, 200, 300) to assess expected model performance across a large series of independent and identically conducted empirical studies. All parameter values were fixed during simulation to assess model performance for social plasticity, assortment, and selection of modest effect size (Pearson r = ±0.3 and Cohen's d = ±0.3; see **Appendix S1** for further details). General-purpose weakly informative priors were used for all model parameters to enhance parameter identification and reduce the risk of inferential bias (Lemoine, 2019). Note that we investigated a sufficiently plausible but simplistic scenario so as to assess the basic properties of SAMs. We therefore did not consider a variety of biologically pertinent processes that are not of direct relevance to our simulation goals, including, among others, differential mortality risk, divorce rates, sex differences in phenotypic expression, and extra-pair matings, all of which have important effects on fitness in socially monogamous species (Culina, Radersma, & Sheldon, 2015; Jeschke & Kokko, 2008; Petrie & Kempenaers, 1998).

Evolutionary payoffs for aggression SRN intercepts and slopes were assumed to be symmetric between serially monogamous male and female partners, so that a single fitness function characterized selection gradients on each individual in the population (**Figure 2**). We assumed that selection gradients and assortment coefficients were constant across seasons, and the population intercept of fitness υ_0 was fixed to 1 for all simulations to provide an appropriate measure of relative fitness (Lande & Arnold, 1983). Following previous work by Thomson et al. (2018), a basic population structure was simulated to derive a random relatedness matrix **A**

used for generating individual SRNs. A simple sorting procedure was then used to generate assortment between social partners' SRN slopes.

Performance metrics

SAM performance was assessed through estimates of bias, dispersion, and power across datasets, with particular attention to estimates of the population-level interaction coefficient (ψ_1), the SRN slope assortment coefficient ($\beta_{\bar{\psi}'\psi}$), the selection differentials ($s_{\bar{\mu}}$ and $s_{\bar{\psi}}$), and the responses to selection ($\Delta\bar{\mu}$ and $\Delta\bar{\psi}$). The selection differentials and responses to selection integrated posterior uncertainty across multiple parameters, capturing the overall ability of the model to predict and explain adaptive social evolution. Note that responses could not be estimated for the phenotypic SAM due to the absence of genetic information. For each model parameter, we calculated the median value as a measure of the central tendency of the posterior distribution. Parameter bias was calculated by subtracting posterior median estimates from the known parameter values used for simulation of each dataset, and then further dividing this quantity by the known value to express bias relative to the total effect size. For example, a median estimate of 0.24 for a true effect of 0.3 would exhibit a bias of -0.06 and a relative bias of -0.2 or -20%. Median absolute bias < |0.2| was interpreted as desirably low for evolutionary inference with modest sample and effect sizes (i.e. parameter accuracy > 80%). Parameter uncertainty was quantified using the median absolute deviation (MAD) divided by the median estimate, providing a robust measure of relative dispersion comparable to a coefficient of variation (Arachchige, Prendergast, & Staudte, 2019). Median uncertainty ≤ |0.5| thus indicated a central tendency at least 2x larger than the uncertainty of the estimate, which we considered desirable for confident statistical inference. Finally, to estimate power, we calculated the posterior probability supporting an effect in the direction of the true effect, with posterior probabilities closer to 1 indicating stronger support for a known positive or negative effect. A median posterior probability ≥ 0.95 would therefore indicate ≤ 0.05 probability in support of

the incorrect direction of the true effect, which we considered desirably low in keeping with standard conventions.

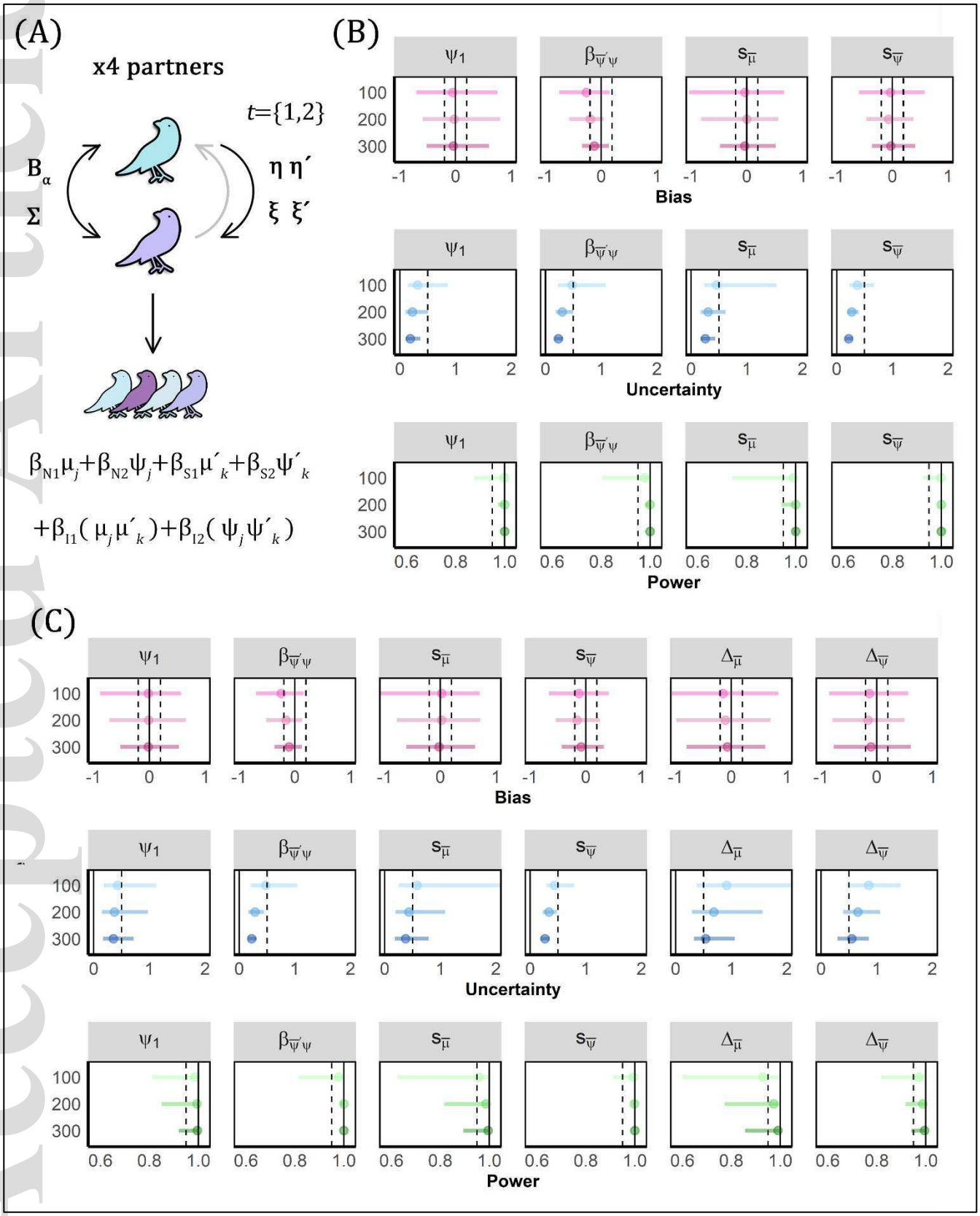
Results

Key simulation results are visualized in **Figure 3**, with results for other parameters summarized in **Appendix S1**. Overall, for both phenotypic (P) and quantitative genetic (QG) analysis, we found that the SAM exhibited desirable low bias, low uncertainty, and high power for modestly sized evolutionary parameters, particularly when N = 200-300. At N = 100, small median biases were observed in the SRN slope assortment coefficient $\beta_{\bar{\psi}'\psi}$ (P: -27%; QG: -24%) and non-social SRN slope selection gradient β_{N2} (P: 21%; QG: 21%), with other model parameters exhibiting desirably low bias (P: -15% - 12%; GQ: -16% - 12%). As explained above, this bias results from the regularization of the model priors, which conservatively pool the assortment coefficients toward zero. However, as expected, median bias steadily decreased with sample size, leading to lower median bias for $\beta_{\bar{\psi}'\psi}$ at N=200 (P: -19%; QG: -16%) and N=300 (P: -12%; GQ: -10%), as well as for the other parameters at N=200 (P: -8% - 17%; QG: -16% - 17%) and N=300 (P: -9% - 10%; QG: -7% - 11%). This means that, for a modestly sized assortment coefficient (equivalent to Pearson r = 0.3), the SAM will be expected to estimate a value of ≈ 0.22 at N = 100 and ≈ 0.26 at N = 200-300.

Parameter uncertainty also steadily decreased across sample sizes. At N=100, median uncertainty was already desirably low for most phenotypic parameters (P: 0.04 – 0.72; QG: 0.04 – 0.61), while genetically influenced parameters exhibited greater median uncertainty (QG: 0.63-0.92). Similar patterns were observed at N=200, with desirably low uncertainty for phenotypic parameters at N = 200 (P: 0.03 – 0.48; QG: 0.03 – 0.48) and greater uncertainty in the genetically influenced parameters (QG: 0.47-0.69). At N=300, all parameters began to exhibit desirably low uncertainty (P: 0.02-0.37; QG: 0.02-0.55).

Posterior probabilities similarly increased across sample sizes. At N = 100, median posterior probabilities were already quite high (P: 0.92 - 1; QG: 0.92 - 1), with even stronger support for true effects found at N = 200 (P: 0.98 - 1; QG: 0.97 - 1) and N = 300 (P: 1; QG: 0.99 - 1). This suggests that both the selection differentials and the direction of genetic response in SRN parameters could be reliably detected at N = 100-200, despite the greater statistical uncertainty observed for the exact magnitude of the genetic parameters. Furthermore, by N = 300, the distribution of posterior probabilities was much narrower across parameters, suggesting that random sampling was less likely to cause inferential error in the direction of evolutionary change, despite its more apparent effects on the distribution of bias and uncertainty across datasets. Similarly, the median posterior probability of $\beta_{\bar{\psi}'\psi}$ was already quite high (QG: 0.98) at N = 100, suggesting that the presence of modest assortment can be reliably detected in smaller samples even if the exact magnitude of assortment will tend to be slightly underestimated. Overall, the results demonstrate that the direction of most effects can be reliably detected at N = 100 with dense individual-level sampling, while sample sizes of N $\geq$ 200 will provide optimal conditions for more accurately and precisely estimating the magnitude of effects.

831 **Figure 3. Simulation results.**



832

Caption. (**A**) A basic overview of the SAM simulation. Each individual was measured twice for aggression t = {1,2} during a breeding season with 4x lifetime mating partners. Associations between individuals and their mates were caused by assortment ($\mathbf{B}_\alpha$) and unmeasured environmental effects (**Σ**; e.g. spatiotemporal heterogeneity), as well as social feedback due to aggression SRNs (**η, η′**) and residual feedback causing further SRN measurement error (**ξ, ξ′**; e.g. unmeasured trait interactions). At the end of each season, breeding success was determined by non-social ($\boldsymbol{\beta}_\mathbf{N}$) and social selection ($\boldsymbol{\beta}_\mathbf{S}$, $\boldsymbol{\beta}_\mathbf{I}$) on the SRN parameters of individuals and their partners. (**A**) Results from the phenotypic SAM for N = 100-300 (y-axis). Results are shown for the bias, uncertainty, and power (posterior probability) of key evolutionary parameters, excluding genetic responses (Δ) due to the absence of genetic information. Regions between the dashed and solid lines indicate desirable model performance, i.e. relative bias < 0.2, uncertainty < 0.5, and power ≥ 0.95. Results across datasets are summarized by median estimates (dot) and 90% CIs (bars) capturing the highest continuous density interval across 200 simulated datasets. (**C**) Results from the quantitative genetic SAM for N = 100-300.

Conclusion

Social interactions play a key role in the evolution of complex phenotypes and the emergence of novel levels of biological organization (Bourke, 2011; Rubenstein & Abbot, 2017; West et al., 2015). Evolutionary quantitative geneticists have developed a large body of theory for predicting the response to selection on interacting phenotypes, as well as for disentangling the individual and social determinants of phenotypic expression (Bijma & Wade, 2008; McGlothlin & Brodie, 2009; McGlothlin et al., 2010; Moore, Brodie, & Wolf, 1997; Wolf et al., 1999). However, despite extensive formal elaboration and a growing body of empirical applications (e.g., Farine & Sheldon, 2015; Fisher et al., 2019; Formica et al., 2011; Santostefano et al., 2017), it remains difficult to specify appropriate trait-based models for disentangling the effects of plasticity, assortment, and selection in the wild, limiting our causal understanding of social evolution. To address this issue, we have proposed social animal models (SAMs) extending the classical animal model to account for the role of social reaction

norms (SRNs) in mediating the repeatable effects of social interactions, as well as to address key statistical challenges in the empirical study of social phenotypes (**Figure 1**).

To demonstrate the empirical application of SAMs and investigate their statistical performance, we simulated data on aggressive interactions and their fitness consequences for assortatively mated avian breeding pairs (**Figure 2**). With sample sizes applicable to long-term field studies, we observed desirably low bias and uncertainty as well as desirably high power for key evolutionary parameters. These models not only detected and rather accurately recovered the magnitude of social plasticity, assortment and selection on SRNs, but also the selection differentials and genetic response in SRNs caused by selection. These results indicate that SAMs provide an integrative and robust approach for investigating adaptive social evolution in the wild. Furthermore, although SAMs are more complex in specification than traditional trait-based models, we have provided extensive coding tutorials for helping researchers to extend these models to their own datasets (**Appendix S1**).

It is important to further emphasize that our simulated system is quite simple, lacking many of the features typical of empirical populations (e.g. spatiotemporal autocorrelation, parental effects, non-linear SRNs, extra-pair matings). Thus, while our results provide a proof-of-principle demonstration that SAMs are robust Bayesian estimators, researchers should be cautious in generalizing our results to more complex model structures, as accurate estimation and detection of many additional random or fixed effects will likely require larger sample sizes. Instead, we encourage others to modify our simulation code to assess the performance of an appropriate SAM relevant to their investigation. In general, fake-data simulation is crucial for ensuring that statistical models have been appropriately specified, as well for benchmarking expected performance prior to data analysis (Gelman, Hill, & Vehtari, 2020). Larger sample sizes will also be required for reliably estimating the effects of multiple phenotypic interactions on trait expression and fitness, each of which may reflect its own multidimensional SRN. In

general, higher dimensionality will rapidly increase the data requirements of any quantitative genetic model (Dochtermann & Roff, 2010), including SAMs. Therefore, it may often be advantageous to use dimension reduction techniques, such as structural equation or generalized network modelling, to further simplify the structure of these integrated SRNs (Araya-Ajoy & Dingemanse, 2014; Martin et al., 2019). Despite these caveats, our results clearly show the desirable performance of quantitative genetic SAMs with relatively modest sample sizes. We have also shown that phenotypic SAMs can readily estimate social plasticity, assortment, and selection even when sufficient genetic information is unavailable. We thus expect that SAMs, employed within a fully Bayesian statistical framework, will readily enhance both phenotypic and quantitative genetic studies of social interactions in the wild.

Conflict of Interest

The authors have no conflict of interest to declare.

References

Alger, I., Hooper, P. L., Cox, D., Stieglitz, J., & Kaplan, H. S. (2020). Paternal provisioning results from ecological change. Proceedings of the National Academy of Sciences, 117, 10746-10754.

Arachchige, C. N., Prendergast, L. A., & Staudte, R. G. (2020). Robust analogs to the coefficient of variation. Journal of Applied Statistics, 1-23.

Araya-Ajoy, Y. G., & Dingemanse, N. J. (2014). Characterizing behavioural 'characters': an evolutionary framework. Proceedings of the Royal Society B, 281, 20132645.

Araya-Ajoy, Y. G., Westneat, D. F., & Wright, J. (2020). Pathways to social evolution and their evolutionary feedbacks. Evolution, 74, 1894-1907.

Bailey, N. W., Marie-Orleach, L., & Moore, A. J. (2018). Indirect genetic effects in behavioral ecology: does behavior play a special role in evolution?. Behavioral Ecology, 29, 1-11.

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. Animal Behaviour, 77, 771-783.

Bijma, P. (2011). A general definition of the heritable variation that determines the potential of a population to respond to selection. Genetics, 189, 1347-1359.

Bijma, P. (2014). The quantitative genetics of indirect genetic effects: a selective review of modelling issues. Heredity, 112, 61-69.

Bijma, P., & Wade, M. J. (2008). The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. Journal of Evolutionary Biology, 21, 1175-1188.

Bollen, K. A., & Noble, M. D. (2011). Structural equation models and the quantification of behavior. Proceedings of the National Academy of Sciences, 108, 15639-15646.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). Time series analysis: forecasting and control (5th edition). John Wiley & Sons.

Bourke, A. F. (2011). Principles of social evolution. Oxford University Press.

Brask, J. B., Croft, D. P., Edenbrow, M., James, R., Bleakley, B. H., Ramnarine, I. W., ... & Darden, S. K. (2019). Evolution of non-kin cooperation: social assortment by cooperative phenotype in guppies. Royal Society Open Science, 6, 181493.

Brommer, J. E. (2013). On between-individual and residual (co) variances in the study of animal personality: are you willing to take the "individual gambit"?. Behavioral Ecology and Sociobiology, 67, 1027-1032.

Cally, J. G., Stuart-Fox, D., & Holman, L. (2019). Meta-analytic evidence that sexual selection improves population fitness. Nature Communications, 10, 1-10.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. Journal of Statistical Software, 76.

948 Carter, A. J., Lee, A. E., Marshall, H. H., Ticó, M. T., & Cowlishaw, G. (2015). Phenotypic
949 assortment in wild primate networks: implications for the dissemination of
950 information. Royal Society Open Science, 2, 140444.

951 Cauchoix, M., Chow, P. K. Y., Van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason,
952 G., ... & Cauchard, L. (2018). The repeatability of cognitive performance: a meta-
953 analysis. Philosophical Transactions of the Royal Society B, 373, 20170281.

954 Chenoweth, S. F., Rundle, H. D., & Blows, M. W. (2010). Experimental evidence for the
955 evolution of indirect genetic effects: changes in the interaction effect coefficient, psi
956 (ψ), due to sexual selection. Evolution, 64, 1849-1856.

957 Class, B., Dingemanse, N. J., Araya-Ajoy, Y. G., & Brommer, J. E. (2017). A statistical
958 methodology for estimating assortative mating for phenotypic traits that are labile or
959 measured with error. Methods in Ecology and Evolution, 8, 1910-1919.

960 Culina, A., Radersma, R., & Sheldon, B. C. (2015). Trading up: the fitness consequences of
961 divorce in monogamous birds. Biological Reviews, 90, 1015-1034.

962 Díaz-Muñoz, S. L., DuVal, E. H., Krakauer, A. H., & Lacey, E. A. (2014). Cooperating to
963 compete: altruism, sexual selection and causes of male reproductive
964 cooperation. Animal Behaviour, 88, 67-78.

965 Dingemanse, N. J., & Araya-Ajoy, Y. G. (2015). Interacting personalities: behavioural ecology
966 meets quantitative genetics. Trends in Ecology & Evolution, 30, 88-97.

967 Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in
968 behaviour: mixed-effect modelling approaches. Journal of Animal Ecology, 82, 39-54.

969 Dingemanse, N. J., Kazem, A. J., Réale, D., & Wright, J. (2010). Behavioural reaction norms:
970 animal personality meets individual plasticity. Trends in Ecology & Evolution, 25, 81-
971 89.

972 Dochtermann, N. A., & Roff, D. A. (2010). Applying a quantitative genetics framework to
973 behavioural syndrome research. Philosophical Transactions of the Royal Society
974 B, 365, 4013-4020.

975 Evans, S. R., Waldvogel, D., Vasiljevic, N., & Postma, E. (2018). Heritable spouse effects
976 increase evolutionary potential of human reproductive timing. Proceedings of the Royal
977 Society B, 285, 20172763.

978 Fanson, K. V., & Biro, P. A. (2019). Meta-analytic insights into factors influencing the
979 repeatability of hormone levels in agricultural, ecological, and medical
980 fields. American Journal of Physiology-Regulatory, Integrative and Comparative
981 Physiology, 316, R101-R109.

982 Farine, D. R., & Sheldon, B. C. (2015). Selection for territory acquisition is modulated by
983 social network structure in a wild songbird. Journal of Evolutionary Biology, 28, 547-
984 556.

985 Fisher, D. N., Wilson, A. J., Boutin, S., Dantzer, B., Lane, J. E., Coltman, D. W., ... & McAdam,
986 A. G. (2019). Social effects of territorial neighbours on the timing of spring breeding
987 in North American red squirrels. Journal of Evolutionary Biology, 32, 559-571.

988 Frank, S. A. (1998). Foundations of social evolution. Princeton University Press.

989 Frank, S. A. (2007). All of life is social. Current Biology, 17, R648-R650.

990 Gabriel, P. O., & Black, J. M. (2012). Behavioural syndromes, partner compatibility and
991 reproductive performance in Steller's jays. Ethology, 118, 76-86.

992 Gelman, A., Hill, J., & Vehtari, A. (2020). Regression and other stories. Cambridge University
993 Press.

994 Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J., & Thompson, R. (2002). ASReml
995 user guide release 1.0.

996 Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models:
997 the MCMCglmm R package. Journal of Statistical Software, 33, 1-22.

998 Hadfield, J. D., & Thomson, C. E. (2017). Interpreting selection when individuals
999 interact. Methods in Ecology and Evolution, 8, 688-699.

1000 Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., & Kruuk, L. E. (2010). The misuse
1001 of BLUP in ecology and evolution. The American Naturalist, 175, 116-125.

1002 Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. Journal of Theoretical
1003 Biology, 7, 17-52.

1004 Hare, R. M., & Simmons, L. W. (2019). Sexual selection and its evolutionary consequences in
1005 female animals. Biological Reviews, 94, 929-956.

1006 Heckerman, D., Gurdasani, D., Kadie, C., Pomilla, C., Carstensen, T., Martin, H., ... &
1007 Kaleebu, P. (2016). Linear mixed model for heritability estimation that explicitly
1008 addresses environmental variation. Proceedings of the National Academy of
1009 Sciences, 113, 7377-7382.

1010 Houslay, T. M., & Wilson, A. J. (2017). Avoiding the misuse of BLUP in behavioural
1011 ecology. Behavioral Ecology, 28, 948-952.

1012 Hughes, W. O., Oldroyd, B. P., Beekman, M., & Ratnieks, F. L. (2008). Ancestral monogamy
1013 shows kin selection is key to the evolution of eusociality. Science, 320, 1213-1216.

1014 Janicke, T., Marie-Orleach, L., Aubier, T. G., Perrier, C., & Morrow, E. H. (2019). Assortative
1015 mating in animals and its role for speciation. The American Naturalist, 194, 865-875.

1016 Jeschke, J. M., & Kokko, H. (2008). Mortality and other determinants of bird divorce
1017 rate. Behavioral Ecology and Sociobiology, 63, 1-9.

1018 Jiang, Y., Bolnick, D. I., & Kirkpatrick, M. (2013). Assortative mating in animals. The
1019 American Naturalist, 181, E125-E138.

1020 Kazancıoğlu, E., Klug, H., & Alonzo, S. H. (2012). The evolution of social interactions changes
1021 predictions about interacting phenotypes. Evolution, 66, 2056-2064.

Kokko, H., & Johnstone, R. A. (2002). Why is mutual mate choice not the norm? Operational sex ratios, sex roles and the evolution of sexually dimorphic and monomorphic signalling. Philosophical Transactions of the Royal Society of London. Series B, 357, 319-330.

Koski, S. E., (2011). How to measure animal personality and why does it matter? Integrating psychological and biological approaches to animal personality. In M. Inoue-Murayama, S. Kawamura, & A. Weiss (Eds.), From genes to animal behavior: Social structures, personalities, communication by color (pp. 115–136). Tokyo:Springer.

Koster, J., Leckie, G., Miller, A., & Hames, R. (2015). Multilevel modeling analysis of dyadic network data with an application to Ye'kwana food sharing. American Journal of Physical Anthropology, 157, 507-512.

Kruuk, L. E. B., & Hadfield, J. D. (2007). How to separate genetic and environmental causes of similarity between relatives. Journal of Evolutionary Biology, 20, 1890-1903.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution, 33, 402-416.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. Evolution, 37, 1210–1226

Lemoine, N. P. (2019). Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. Oikos, 128, 912-928.

Leniowski, K., & Węgrzyn, E. (2018). Equal division of parental care enhances nestling development in the Blackcap. PloS one, 13, e0207757.

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. Science, 355, 584-585.

Marshall, J. A. (2015). Social evolution and inclusive fitness theory: an introduction. Princeton University Press.

Martin, J. S. (2021). Estimating Nonlinear Selection on Behavioral Reaction Norms. EcoEvoRxiv preprint, doi: 10.32942/osf.io/u26tz

Martin, J. S., Massen, J. J., Šlipogor, V., Bugnyar, T., Jaeggi, A. V., & Koski, S. E. (2019). The EGA+ GNM framework: An integrative approach to modelling behavioural syndromes. Methods in Ecology and Evolution, 10, 245-257.

McAdam, A. G., Garant, D., & Wilson, A. J. (2014). The effects of others' genes: maternal and other indirect genetic effects. In A. Charmantier, D. Garant, & L.E.B. Kruuk (Eds.), Quantitative genetics in the wild (pp. 84-103). Oxford: Oxford University Press.

McCullough, E. L., Miller, C. W., & Emlen, D. J. (2016). Why sexually selected weapons are not ornaments. Trends in Ecology & Evolution, 31, 742-751.

McDonald, G. C., Farine, D. R., Foster, K. R., & Biernaskie, J. M. (2017). Assortment and the analysis of natural selection on social traits. Evolution, 71, 2693-2702.

McElreath, R. (2020). Statistical rethinking: A Bayesian course with examples in R and Stan (2nd Edition). CRC press.

McGlothlin, J. W., Akçay, E., Brodie, E. D., Moore, A. J., & Van Cleve, J. (2021). A synthesis of game theory and quantitative genetic models of social evolution. bioRxiv. doi: https://doi.org/10.1101/2021.03.27.437341.

McGlothlin, J. W., & Brodie III, E. D. (2009). How to measure indirect genetic effects: the congruence of trait-based and variance-partitioning approaches. Evolution, 63, 1785-1795.

McGlothlin, J. W., Moore, A. J., Wolf, J. B., & Brodie III, E. D. (2010). Interacting phenotypes and the evolutionary process. III. Social evolution. Evolution, 64, 2558-2574.

McGlothlin, J. W., Wolf, J. B., Brodie III, E. D., & Moore, A. J. (2014). Quantitative genetic versions of Hamilton's rule with empirical applications. Philosophical Transactions of the Royal Society B, 369, 20130358.

McNamara, J. M., & Leimar, O. 2020. Game Theory in Biology. Oxford, UK: Oxford University Press.

Molleman, L., Van den Berg, P., & Weissing, F. J. (2014). Consistent individual differences in human social learning strategies. Nature Communications, 5, 1-9.

Moore, A. J., Brodie III, E. D., & Wolf, J. B. (1997). Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. Evolution, 51, 1352-1362.

Morrissey, M. B., Kruuk, L. E. B., & Wilson, A. J. (2010). The danger of applying the breeder's equation in observational studies of natural populations. Journal of Evolutionary Biology, 23, 2277-2288.

Morrissey, M. B., Parker, D. J., Korsten, P., Pemberton, J. M., Kruuk, L. E., & Wilson, A. J. (2012). The prediction of adaptive evolution: empirical application of the secondary theorem of selection and comparison to the breeder's equation. Evolution, 66, 2399-2410.

Morrissey, M. B., & Sakrejda, K. (2013). Unification of regression-based methods for the analysis of natural selection. Evolution, 67, 2094-2100.

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews, 85, 935-956.

Nussey, D. H., Wilson, A. J., & Brommer, J. E. (2007). The evolutionary ecology of individual phenotypic plasticity in wild populations. Journal of Evolutionary Biology, 20, 831-844.

Parejo, D., & Avilés, J. M. (2020). Melanism influences the use of social information in a polymorphic owl. Scientific Reports, 10, 1-9.

Petrie, M., & Kempenaers, B. (1998). Extra-pair paternity in birds: explaining variation between species and populations. Trends in Ecology & Evolution, 13, 52-58.

1097 Pilakouta, N., Hanlon, E. J., & Smiseth, P. T. (2018). Biparental care is more than the sum of
1098 its parts: experimental evidence for synergistic effects on offspring fitness. Proceedings
1099 of the Royal Society B, 285, 20180875.

1100 van de Pol, M., and Wright, J. (2009). A simple method for distinguishing within- versus
1101 between-subject effects using mixed models. Animal Behaviour, 77, 753–758.

1102 Price, G. R. (1972). Extension of covariance selection mathematics. Annals of Human
1103 Genetics, 35, 485-490.

1104 Queller, D. C. (2011). Expanded social fitness and Hamilton's rule for kin, kith, and
1105 kind. Proceedings of the National Academy of Sciences, 108, 10792-10799.

1106 R Core Team (2013). R: A language and environment for statistical computing. R Foundation
1107 for Statistical Computing, Vienna, Austria. http://www.**R**-project.org/.

1108 Ramakers, J. J., Gienapp, P., & Visser, M. E. (2019). Phenological mismatch drives selection
1109 on elevation, but not on slope, of breeding time plasticity in a wild
1110 songbird. Evolution, 73, 175-187.

1111 Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. Animal
1112 Science, 8, 95-108.

1113 Rubenstein, D. R., & Abbot, P. (Eds.). (2017). Comparative social evolution. Cambridge
1114 University Press.

1115 Santostefano, F., Wilson, A. J., Niemelä, P. T., & Dingemanse, N. J. (2017). Indirect genetic
1116 effects: a key component of the genetic architecture of behaviour. Scientific Reports, 7,
1117 1-9.

1118 Searle, S. R. (1961). Phenotypic, genetic, and environmental correlations. Biometrics, 17, 474–
1119 480.

1120 Silva, J. C. E., Potts, B. M., Bijma, P., Kerr, R. J., & Pilbeam, D. J. (2013). Genetic control of
1121 interactions among individuals: contrasting outcomes of indirect genetic effects arising
1122 from neighbour disease infection and competition in a forest tree. New
1123 Phytologist, 197, 631-641.

1124 Smith, J. M., & Harper, D. (2003). Animal signals. Oxford University Press.

1125 Sprau, P., & Dingemanse, N. J. (2017). An approach to distinguish between plasticity and non-
1126 random distributions of behavioral types along urban gradients in a wild passerine
1127 bird. Frontiers in Ecology and Evolution, 5, 92.

1128 Steglich, C., Snijders, T. A., & Pearson, M. (2010). Dynamic networks and behavior:
1129 Separating selection from influence. Sociological Methodology, 40, 329-393.

1130 Stinchcombe, J. R., Simonsen, A. K., & Blows, M. W. (2014). Estimating uncertainty in
1131 multivariate responses to selection. Evolution, 68, 1188-1196.

1132 Strickland, K., & Frère, C. H. (2019). Individual variation in the social plasticity of water
1133 dragons. The American Naturalist, 194, 194-206.

Thomson, C. E., Winney, I. S., Salles, O. C., & Pujol, B. (2018). A guide to using a multiple-matrix animal model to disentangle genetic and nongenetic causes of phenotypic variance. PloS one, 13.

Van Cleve, J. (2017). Stags, hawks, and doves: Social evolution theory and individual variation in cooperation. Integrative and Comparative Biology, 57, 566-579.

Van Cleve, J., & Akçay, E. (2014). Pathways to social evolution: reciprocity, relatedness, and synergy. Evolution, 68, 2245-2258.

de Villemereuil, P., Schielzeth, H., Nakagawa, S., & Morrissey, M. (2016). General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics, 204, 1281-1294.

Wade, M. J., Bijma, P., Ellen, E. D., & Muir, W. (2010). Group selection and social evolution in domesticated animals. Evolutionary Applications, 3, 453-465.

Wang, D., Forstmeier, W., Valcu, M., Dingemanse, N. J., Bulla, M., Both, C., ... & Kempenaers, B. (2019). Scrutinizing assortative mating in birds. PLoS Biology, 17, e3000156.

Watson, S. K., Vale, G. L., Hopper, L. M., Dean, L. G., Kendal, R. L., Price, E. E., ... & Whiten, A. (2018). Chimpanzees demonstrate individual differences in social information use. Animal Cognition, 21, 639-650.

Weis, A. E., & Gorman, W. L. (1990). Measuring selection on reaction norms: an exploration of the Eurosta-Solidago system. Evolution, 44, 820-831.

West, S. A., Fisher, R. M., Gardner, A., & Kiers, E. T. (2015). Major evolutionary transitions in individuality. Proceedings of the National Academy of Sciences, 112, 10112-10119.

Westneat, D. F. (2012). Evolution in response to social selection: the importance of interactive effects of traits on fitness. Evolution, 66, 890-895.

Westneat, D. F., Wright, J., & Dingemanse, N. J. (2015). The biology hidden inside residual within-individual phenotypic variation. Biological Reviews, 90, 729-743.

Westneat, D. F., Araya-Ajoy, Y. G., Allegue, H., Class, B., Dingemanse, N., Dochtermann, N. A., ... & Schielzeth, H. (2020). Collision between biological process and statistical analysis revealed by mean centring. Journal of Animal Ecology, 89, 2813-2824.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., ... & Nussey, D. H. (2010). An ecologist's guide to the animal model. Journal of Animal Ecology, 79, 13-26.

Wolf, J. B., Brodie III, E. D., & Moore, A. J. (1999). Interacting phenotypes and the evolutionary process. II. Selection resulting from social interactions. The American Naturalist, 153, 254-266.

Wolf, M., Van Doorn, G. S., & Weissing, F. J. (2008). Evolutionary emergence of responsive and unresponsive personalities. Proceedings of the National Academy of Sciences, 105, 15825-15830.